PROCEDE ET DISPOSITIF FONCTIONNANT EN TEMPS REEL, POUR LE REPERAGE ET LA LOCALISATION D'UNE ZONE EN MOUVEMENT RELATIF DANS UNE SCENE, AINSI QUE POUR LA DETERMINATION DE LA VITESSE ET DE LA DIRECTION DU DEPLACEMENT

La présente invention a pour objet un procédé et un dispositif permettant de repérer et de localiser une zone en mouvement relatif dans une scène et de déterminer la vitesse et la direction orientée de ce mouvement relatif, et ceci en temps réel.

Par mouvement relatif, on entend aussi bien le mouvement de ladite zone (qui peut être constituée par un «objet», au sens le plus large incluant un être vivant ou une portion d'un être vivant, une main par exemple) dans un environnement sensiblement immobile, que l'immobilité plus ou moins complète de ladite zone (ou «objet») dans un environnement en déplacement au moins partiel.

L'invention est relative au traitement d'un signal vidéo numérique en provenance d'un système d'observation, constitué par un système optique d'entrée ou objectif, apte à former une image de la scène observée, et par un système de conversion optoélectronique ou capteur, apte à convertir ladite image qu'il reçoit en un signal numérique de sortie.

En général le système d'observation est constitué par une caméra vidéo ou caméscope, qui observe la scène à surveiller (ledit signal de sortie numérique étant alors constitué par le signal vidéo numérique débité par une caméra à sortie numérique ou par la sortie d'un convertisseur analogique/numérique dont l'entrée est connectée à la sortie d'une caméra débitant un signal vidéo analogique).

Le système d'observation pourrait également être constitué par l'objectif d'un instrument d'optique (jumelles, lunette d'observation, viseur), dont on prélève au moins une portion du faisceau lumineux sortant, et par un capteur photo-électronique, de type CCD ou CMOS par exemple, avec l'électronique associée habituelle, capteur recevant l'image formée par ladite portion de faisceau lumineux et le convertissant, par l'électronique associée, en un signal vidéo numérique de sortie.

L'invention consiste essentiellement à traiter le signal vidéo numérique de sortie d'un système d'observation, notamment d'une caméra vidéo, à sortie numérique pour en déduire des signaux signalant l'existence et la localisation d'une zone en déplacement relatif dans ladite scène, ainsi que la vitesse et la direction orientée du déplacement dans le cas où ladite zone se déplace effectivement dans ladite scène relativement à un environnement sensiblement immobile, et ceci en temps réel.

Le système le plus perfectionné pour repérer et localiser un objet en mouvement relatif et déterminer sa vitesse et sa direction orientée de déplacement est le système de la vision animale ou humaine, par exemple d'un chasseur à l'affût localisant le déplacement d'un animal, ainsi que la direction et la vitesse de ce déplacement.

Dans la technique antérieure on a proposé des dispositifs de surveillance du type rétine artificielle, soit analogiques (Giocomo Indiveri et al. in Proceedings of MicroNeuro'96 p. 15 à 22), soit numériques (Pierre-François Rüedi in Proceedings of MicroNeuro'96 p. 23 à 29), mais il s'agit dans le premier article de détecteurs et unités analogiques à structure complexe et dans le second article de moyens de repérage des bords d'un objet ; en outre dans les dispositifs décrits on a recours à des mémoires très rapides et de grande capacité pour pouvoir opérer en temps réel, et on obtient des renseignements limités en ce qui concerne les zones ou objets en mouvement.

On a ainsi proposé de mémoriser, dans une première mémoire bi-dimensionnelle, le signal d'une trame en provenance d'une caméra vidéo, ou analogue, constitué par une suite de données concernant les pixels représentatifs de la scène observée par la caméra à un instant t_0, puis, dans une deuxième mémoire bi-dimensionnelle, le signal vidéo, pour la trame correspondante suivante, représentatif de ladite scène à un instant t_1. Si un objet s'est déplacé entre t_0 et t_1, on détermine, d'une part, la distance d parcourue par celui-ci dans la scène entre t_1 et t_0 et, d'autre part, la durée $T = t_1 - t_0$ entre les débuts de deux trames correspondantes successives relatives aux mêmes pixels. La vitesse du déplacement est alors égale à d/T. Un tel système nécessite une capacité totale de mémoire très importante si on désire obtenir des indications précises de vitesse et de direction orientée caractérisant le déplacement. En outre, un certain retard existe en ce qui concerne l'obtention des indications de vitesse et de direction du déplacement ; en effet de telles informations ne sont disponibles qu'à l'instant $t_1 + R$, en appelant R la durée des calculs portant sur l'intervalle $t_0 - t_1$. Ce double inconvénient (nécessité d'une grande capacité de mémoire et retard à l'obtention des informations désirées) limite les applications d'un tel système.

Par ailleurs le brevet français No 2.611.063, dont l'un des inventeurs (Monsieur Patrick Pirim) est l'inventeur de la présente invention, décrit un procédé et un dispositif de traitement en temps réel d'un flot de données séquencé, constitué en particulier par le signal de sortie d'un caméscope, afin de réaliser une compression des données. Selon ce brevet antérieur, on forme l'histogramme des niveaux du signal suivant une loi de classification pour une première séquence, on mémorise la fonction de Gauss représentative associée à cet histogramme, dont on extrait les niveaux maximum et minimum, on compare les niveaux de la séquence ultérieure, ou deuxième séquence, aux dits niveaux du signal pour la première séquence, mémorisé avec une constante de temps constante, identique pour chaque pixel, on engendre un signal binaire de classification qui

caractérise ladite séquence suivante par rapport à la loi de classification, on engendre, à partir de ce signal binaire, un signal auxiliaire représentatif de la durée et de la position d'une plage de valeurs significatives et enfin on engendre, à partir dudit signal auxiliaire, un signal de localisation de la plage ayant la plus longue durée, dite plage dominante; et on répète ces opérations pour les séquences suivantes du signal séquencé. Ce procédé et ce dispositif de classification permettent une compression des données en ne retenant que les paramètres intéressants du flot traité de données séquencé. En particulier ce procédé permet de traiter un signal vidéo numérique représentatif d'une image vidéo en vue d'extraire et localiser au moins une caractéristique d'au moins une zone de ladite image. On peut ainsi classer les niveaux de luminance et/ou de chrominance du signal et caractériser et localiser un objet dans l'image.

Quant au brevet des Etats-Unis n° 5 488 430, il réalise la détection et l'estimation d'un déplacement en déterminant séparément les changements horizontaux et verticaux de l'image de la zone observée. On y utilise des signaux de différence pour détecter des déplacements de la droite vers la gauche, ou inversement et du haut vers le bas, ou inversement, dans les directions horizontale et verticale respectivement, en effectuant la fonction logique OU EXCLUSIF sur des signaux de différence horizontaux/verticaux et des signaux de différence de trames, d'une part, et en utilisant un rapport des sommes de signaux horizontaux/verticaux et des sommes des signaux de différence de trames par rapport à une fenêtre K x 3, d'autre part. Dans ce brevet U.S. 5 488 430 on utilise les valeurs calculées de l'image suivant les deux directions orthogonales horizontale et verticale avec un écart répétitif identique K dans ces deux directions orthogonales, cet écart K étant défini en fonction des vitesses de déplacement qu'on cherche à déterminer. Le dispositif selon ce brevet U.S. détermine la direction des mouvements suivant chacune des deux directions orthogonales en appliquant aux signaux de différence un ensemble d'opérations de calcul indiquées aux colonnes 12 (en début et en fin) et 13 (en début) qui nécessite des opérateurs électroniques, notamment de division, de multiplication et de sommation, fort complexes (donc difficiles à réaliser) ; des opérateurs complexes supplémentaires sont en outre nécessaires pour obtenir, à partir des projections sur les deux axes horizontal et vertical, la vitesse et la direction orientée du déplacement (extraction de racine carrée pour obtenir l'amplitude de la vitesse et calcul de la fonction arctg pour obtenir la direction orientée). On ne prévoit pas, enfin, dans le brevet 5 488 430 la mise en oeuvre d'un lissage des valeurs de pixel au moyen d'une constante de temps, variable pour chaque pixel, afin de compenser les variations trop rapides de ces valeurs.

Au contraire, la mise en oeuvre du procédé selon l'invention est réalisée au moyen d'un dispositif, faisant l'objet de l'invention, qui est de type numérique, avec une structure relativement simple et une mémoire de capacité relativement réduite, et permet l'obtention rapide des

informations désirées, avec des résultats et des applications très variées (qui correspondent suivant l'application à une demi-image complète dans le cas de trames entrelacées ou à une image complète).

Un article par Alberto Tomita Sr. et Rokuya Ishii, intitulé «Hand Shape Extraction from a Sequence of Digitized Gray-Scale Images», dans Institute of Electrical and Electronics Engineers, vol. 3, 1994, p. 1925-1930, met en oeuvre une détection de mouvement par soustraction entre images successives, suivie de la formation d'histogrammes basée sur la forme de la main humaine, afin d'extraire la forme d'une main humaine dans une scène numérisée. L'analyse par histogramme est basée sur une échelle de gris inhérente à la main humaine. On ne prévoit aucune formation d'histogrammes de coordonnées du plan. Le but unique des auteurs de cet article est de détecter les déplacements d'une main humaine, par exemple afin de remplacer, pour l'introduction de données dans un ordinateur, la souris habituelle par une main dont on repère les déplacements.

Au contraire la présente invention n'est pas limitée à détecter le déplacement d'une main, mais permet de détecter le déplacement relatif de tout objet, au sens le plus large, dans une scène et n'utilise pas des histogrammes basés sur les valeurs de gris d'une main, mais des histogrammes de certaines variables numériques particulières représentatives du déplacement éventuel et des histogrammes de coordonnées du plan.

Selon la présente invention :

- on traite un signal vidéo numérique, en provenance d'un système d'observation, signal constitué, à la manière connue, par une succession de trames (qui correspondent à une demi-image dans le cas de deux trames entrelacées par image ou à une image complète dans le cas d'une seule trame par image) comportant chacune un nombre déterminé de lignes successives et, dans chacune de ces lignes, un nombre déterminé de pixels ou points d'image,
- afin d'obtenir, en utilisant une mémoire de capacité relativement faible, des signaux aptes à indiquer s'il existe, dans la scène observée, une zone en déplacement relatif et, dans ce cas, à préciser la localisation, ainsi que la vitesse et la direction (orientée) de ladite zone si elle se déplace effectivement par rapport à son environnement,
- et ceci en élaborant deux signaux numériques caractéristiques, l'un, d'une variation ou non-variation significative du signal de pixel pour le même emplacement de pixel entre deux trames correspondantes successives et, l'autre, de l'amplitude de cette variation, lorsqu'elle existe, et en répartissant matriciellement ces deux signaux pour les pixels d'une portion de trame à un même instant.

L'invention a tout d'abord pour objet un procédé, en temps réel, pour le repérage et la localisation d'une zone en mouvement relatif dans une scène observée par un système

d'observation à sortie constituée par un signal vidéo numérique du type comportant une succession de trames correspondantes, composées chacune d'une succession de lignes, composées chacune d'une succession de pixels, ainsi que pour la détermination de la vitesse et de la direction orientée du déplacement, ledit procédé étant caractérisé en ce qu'il consiste à effectuer sur le signal vidéo numérique de sortie successivement :

- un traitement de lissage dudit signal vidéo numérique de sortie mettant en oeuvre une constante de temps numérique dont la valeur numérique peut être modifiée pour chacun des pixels dudit signal de sortie, indépendamment pour chacun d'entre eux ;
- une mise en mémoire d'une trame dudit signal de sortie après le lissage, d'une part, et de la constante de temps de lissage associée à la dite trame, d'autre part ;
- un traitement temporel consistant, pour chaque position de pixel, à déterminer l'existence, d'une part, et l'amplitude, d'autre part, d'une variation significative de l'amplitude du signal de pixel entre la trame actuelle et la trame juste antérieure lissée, mise en mémoire, et à générer deux signaux numériques, le premier signal étant un signal binaire ou monobit à deux valeurs possibles dont l'une représente l'existence d'une telle variation significative et l'autre l'absence d'une telle variation entre deux trames successives, la valeur dudit signal binaire modifiant la valeur mémorisée de ladite constante de temps afin de la diminuer si ledit signal représente une variation significative et afin de l'augmenter si ce signal ne représente pas une telle variation , la diminution ou l'augmentation étant réalisée d'une manière quantifiée, tandis que le second signal numérique, dit d'amplitude, est un signal multibits à nombre limité de bits, quantifiant l'amplitude de cette variation ; et
- un traitement spatial consistant, pour chaque trame de signal vidéo numérique d'entrée,
  - à répartir, afin de caractériser les valeurs des pixels, en une matrice à nombre de lignes et nombre de colonnes réduits, par rapport respectivement au nombre de lignes et au nombre de pixels par ligne dans le signal vidéo, uniquement les valeurs, au même instant d'observation, pour une fraction des pixels d'une trame - fraction qui défile par balayage à travers ladite matrice pendant la durée d'une trame - dudit signal binaire, d'une part, et dudit signal numérique d'amplitude, d'autre part,
  - à déterminer, dans cette double représentation matricielle instantanée, une zone particularisée, dans laquelle à la fois ledit signal binaire a la valeur recherchée représentant la présence, ou respectivement l'absence, de variation significative et ledit signal numérique d'amplitude varie, ou respectivement ne varie pas, d'une valeur significative entre pixels voisins dans la matrice suivant une direction orientée à partir d'un pixel d'origine, et ceci pour une même portion de trame, donc à un même instant d'observation, et

- à engendrer des signaux représentatifs de l'existence et de la localisation de la zone en déplacement relatif, ainsi que la vitesse relative intertrames et la direction orientée de ce déplacement, s'il existe, relativement à son environnement, à partir de la répartition matricielle instantanée de ces deux signaux numériques, binaire et d'amplitude.

De préférence, le procédé selon invention est caractérisé en ce qu'il consiste, en outre,

- à former les histogrammes des valeurs des signaux répartis matriciellement, d'une part, et les histogrammes des inclinaisons de deux axes, à pente variable, de coordonnées dans un plan, d'autre part,
- à repérer, dans chaque histogramme formé, un domaine de variation significative de la valeur traitée et
- à déduire, de chaque domaine repéré, l'existence et la localisation, ainsi que la vitesse et la direction orientée, d'une zone en mouvement relatif.

Dans des modes de réalisation particuliers :

- ladite matrice est une matrice carrée à même nombre impair $(2l + 1)$ de lignes et de colonnes, et on considère les matrices emboîtées de 3 x 3, 5 x 5, 7 x 7,...$(2l + 1)$ x $(2l + 1)$ éléments centrées sur le centre de cette matrice carrée afin de déterminer la matrice emboîtée de plus petite taille dans laquelle ledit signal numérique varie dans une direction orientée à partir dudit centre, la valeur dudit signal binaire représentant un dépassement du seuil selon cette direction,
- ladite matrice est une matrice hexagonale et on considère les matrices hexagonales emboîtées de taille croissante centrées sur le centre de cette matrice hexagonale afin de déterminer la matrice emboîtée de plus petite taille dans laquelle ledit signal numérique varie dans une direction orientée,
- ladite matrice est une matrice en L renversé à une seule ligne et une seule colonne et on considère les matrices emboîtées de 3 x 3 pixels, 5 x 5 pixels, 7 x 7 pixels... $(2l + 1)$ x $(2l + 1)$ pixels, pour la ligne et la colonne uniques, afin de déterminer la matrice de plus petite taille dans laquelle le signal varie dans une direction orientée, à savoir la ligne de plus grande pente à quantification constante.

Avantageusement ladite constante de temps est de la forme 2^p, p étant un nombre inférieur à 16, pouvant donc être exprimé par pas plus de 4 bits, la diminution ou l'augmentation de la constante de temps étant réalisée par la soustraction ou l'addition d'une unité à p.

Dans ce cas, si on le désire, on considère des portions successives décroissantes de trames complètes suivant l'algorithme temps - échelle de Mallat et sélectionne la plus grande de ces portions, qui donne des indications de déplacement, vitesse et orientation, compatible avec la valeur de p.

L'invention a également pour objet un dispositif, fonctionnant en temps réel, pour le repérage et la localisation d'une zone en mouvement relatif dans une scène observée par un système d'observation à sortie constituée par un signal vidéo numérique, du type comportant une succession de trames correspondantes, des lignes successives dans chaque trame correspondante et des pixels successifs dans chaque ligne, ainsi que pour la détermination de la vitesse et de la direction orientée du déplacement, par mise en oeuvre du procédé susmentionné, ledit dispositif recevant, en entrée, ledit signal vidéo de sortie et étant caractérisé en ce qu'il comporte en combinaison :

- des moyens de lissage dudit signal vidéo numérique de sortie mettant en oeuvre une constante de temps numérique dont la valeur numérique peut être modifiée pour chacun des pixels dudit signal de sortie, indépendamment pour chacun d'entre eux ;
- des moyens de mémorisation d'une trame dudit signal de sortie après lissage, d'une part, et de la constante de temps de lissage associée à ladite trame, d'autre part ;
- une unité de traitement temporel pour analyser les variations temporelles de l'amplitude du signal de pixel, pour une même position de pixel, entre la trame actuelle et la trame juste antérieure lissée, mise en mémoire, dudit signal vidéo numérique, ladite unité comportant, en association avec une mémoire apte à recevoir, stocker et restituer des informations relatives à la trame correspondante précédente lissée, des moyens de comparaison pour déterminer, pour chaque position de pixel dans la trame du signal vidéo entrant, si la valeur absolue de la différence entre le signal de pixel actuel et une valeur représentative du signal de pixel, pour la même position de pixel, dans la trame antérieure, valeur représentative stockée dans ladite mémoire, dépasse ou non un seuil, en générant un signal binaire ou monobit à deux valeurs, dont l'une représente l'existence d'un dépassement et dont l'autre représente l'absence d'un dépassement, et des moyens de calcul aptes à déterminer un signal numérique d'amplitude multibits, à nombre réduit de bits, dont la valeur est fonction de l'amplitude de la variation de la valeur du même pixel entre la trame actuelle et la trame juste antérieure lissée, mise en mémoire, du signal vidéo numérique ; et
- un ensemble de traitement spatial, dont les entrées reçoivent, de l'unité de traitement temporel, lesdits signaux binaire et numérique d'amplitude successifs pour les pixels d'une même trame, ladite unité comportant des moyens aptes à caractériser les valeurs d'amplitude des pixels, ces moyens répartissant, suivant une matrice à nombre de lignes et nombre de colonnes réduits par rapport respectivement au nombre de lignes et au nombre de pixels par ligne dans une trame dudit signal vidéo numérique, uniquement lesdits signaux binaire et numérique d'amplitude relatif à un même instant, c'est-à-dire à une même trame, celle-ci défilant par balayage à travers

ladite matrice pendant la durée d'une trame, des moyens de repérage pour déterminer, dans ladite matrice, une zone de pixels dans laquelle, à cet instant, le signal binaire a la valeur recherchée et des moyens pour déterminer, dans ladite matrice, une zone de pixels dans laquelle, à ce même instant, le signal numérique d'amplitude varie d'une quantité significative entre pixels voisins, et des moyens qui, en réponse aux indications des deux derniers moyens précédents, génèrent des signaux représentatifs de cette zone de pixels, donc de l'existence et de la localisation d'une zone en mouvement relatif dans la scène observée, ainsi que de la vitesse relative intertrames et de la direction orientée de cette zone lorsqu'elle se déplace effectivement relativement à son environnement.

De préférence l'unité de traitement spatial comporte, en outre, une sortie débitant un signal vidéo numérique retardé constitué par le signal vidéo numérique d'entrée retardé d'une durée égale à la durée des lignes d'une matrice diminuée de la durée d'une ligne afin de fournir un signal de sortie contemporain de l'analyse de la matrice dans ladite unité de traitement temporel.

De préférence, dans le dispositif selon l'invention, l'ensemble de traitement spatial comporte des moyens de retard en cascade dont chacun impose un retard égal à l'écart temporel entre les débuts de deux lignes successives et des moyens à retard en cascade pour chaque ligne imposant chacun un retard égal à l'écart temporel entre deux pixels successifs d'une ligne, les sorties de l'ensemble de tous les registres et de l'entrée des premiers registres de chaque ligne débitant à un instant donné les valeurs dudit signal binaire et dudit signal numérique d'amplitude, à un même instant, vers lesdits moyens de repérage.

Avantageusement le dispositif selon l'invention comporte également des moyens formant les histogrammes des valeurs de sortie de ladite unité de traitement spatial, ainsi que les histogrammes des inclinaisons de deux axes, à pente variable, de coordonnées dans un plan, des moyens pour repérer, dans chaque histogramme, un domaine de variation significative de la valeur traitée, afin de valider ce domaine en leur sortie et de déduire, pour l'ensemble des histogrammes, des signaux de sortie qui signalent et localisent une zone de la scène observée en mouvement relatif, si elle existe, ainsi que la vitesse et la direction orientée de ce mouvement si ladite zone se déplace effectivement par rapport à son environnement.

Si on désire détecter le mouvement d'un objet dans un environnement sensiblement immobile, on détermine la zone de la matrice dans laquelle simultanément le signal binaire a la valeur correspondant à un dépassement du seuil et le signal numérique d'amplitude varie d'une valeur significative entre pixels voisins d'une trame.

Au contraire, si on désire détecter l'immobilité d'un objet dans un environnement sensiblement en mouvement, on détermine la zone de la matrice dans laquelle simultanément le

signal binaire a la valeur correspondant à un non-dépassement du seuil et le signal numérique d'amplitude ne varie pas entre pixels voisins d'une trame.

De préférence, dans le dispositif pour le repérage, la localisation et la détermination de la vitesse et de la direction orientée du déplacement d'une zone en mouvement relatif dans une scène, par mise en oeuvre du procédé susmentionné :

- lesdits moyens de lissage comportent une entrée qui reçoit ledit signal vidéo numérique et calculent, pour les pixels successifs d'une trame de ce signal vidéo, un signal lissé, dans lequel les variations temporelles du signal vidéo numérique d'entrée sont diminuées, par mise en oeuvre d'un signal de seuil reçu sur une autre entrée et d'une constante de temps relative à chaque position de pixel d'une trame, dont la valeur est successivement modifiée afin que le lissage conserve, tout en la réduisant, la tendance de variation du signal vidéo numérique entrant, ces moyens de lissage coopérant avec l'unité de mémoire qui reçoit, stocke et restitue les valeurs actualisées, pour chaque portion de pixel d'une trame, du signal lissé et de ladite constante de temps et débitent, sur leurs sorties, au moins la succession, pour chaque position de pixel, des valeurs de la constante de temps actualisée et des valeurs d'un signal binaire de dépassement ou non-dépassement dudit seuil par la valeur absolue de la différence entre la valeur du pixel et sa valeur lissée ; ladite unité de traitement spatial réalisant la répartition matricielle, par lignes et colonnes en nombre réduit, des sorties desdits moyens de lissage, à savoir des valeurs successives de la constante de temps et dudit signal binaire ;
- on prévoit des moyens de repérage pour repérer, dans ladite répartition matricielle, une zone de pixels dans laquelle, à la fois, soit la valeur dudit signal binaire correspond à un dépassement de seuil et ladite constante de temps varie entre pixels voisins d'une valeur significative dans une direction, et pour produire des signaux de sortie indiquant la localisation de ladite zone et la vitesse et la direction orientée du déplacement dans ladite zone, soit la valeur dudit signal binaire correspond à un non-dépassement du seuil et ladite constante de temps ne varie pas entre pixels voisins et pour produire des signaux de sortie indiquant la localisation de ladite zone.

Dans les modes de réalisation préférés :

- lesdits moyens de lissage comprennent, en combinaison avec une mémoire vidéo ou mémoire de trame (field memory) qui stocke les valeurs successives, pour chaque pixel de trame, de ladite constante de temps et du signal vidéo numérique lissé, des moyens de calcul, pour chaque pixel, de la valeur absolue de la différence entre la valeur du signal vidéo numérique provenant de la caméra et la valeur du signal vidéo numérique précédent lissé, des moyens pour comparer cette différence à un seuil et pour engendrer un signal binaire, dont une des

deux valeurs indique un dépassement dudit seuil et l'autre valeur indique un non-dépassement de celui-ci ; des moyens de mise à jour de la constante de temps recevant de ladite mémoire la valeur juste antérieure de la constante de temps et la diminuant si elle reçoit un signal binaire dont la valeur indique un dépassement de seuil, mais l'augmentant si la valeur du signal binaire indique un non-dépassement, la diminution ou l'augmentation n'ayant toutefois pas lieu si elle aboutissait respectivement à une valeur négative ou à une valeur supérieure à une valeur seuil, et des moyens de mise à jour de la valeur lissée du signal vidéo numérique qui ajoutent algébriquement à la valeur antérieure de ce signal lissé reçue de ladite mémoire le quotient de la différence entre le signal vidéo numérique en provenance de la camera et le signal vidéo numérique lissé précédent en provenance de ladite mémoire par un facteur égal à la valeur de la constante de temps antérieure en provenance de ladite mémoire ;

- ladite constante de temps est sous la forme 2^p, p étant un nombre entier inférieur à 16, pouvant donc être représenté par au plus 4 bits, la diminution ou l'augmentation de la constante de temps étant réalisée par la soustraction ou l'addition d'une unité à p ;
- lesdits moyens de répartition matricielle comprennent, en combinaison avec des moyens de retard appliquant des retards successifs, égaux à la durée d'une ligne du signal vidéo, au signal vidéo numérique en provenance de la caméra pour débiter sur une succession des sorties, ce signal vidéo retardé de la durée d'un retard, de deux retards et ainsi de suite jusqu'à un nombre de retards égal au nombre de lignes de la matrice de répartition spatiale diminué d'une unité, des moyens de répartition matricielle suivant les lignes successives de la matrice recevant, d'une part, le signal vidéo numérique de la caméra non retardé et ce signal progressivement retardé en provenance des sorties des moyens de retard et, d'autre part, les valeurs de la constante de temps et dudit signal binaire en provenance desdits moyens de lissage, pour réaliser une répartition matricielle, par lignes et par colonnes, à un instant donné desdites valeurs de la constante de temps et du signal binaire pour les pixels d'une portion de trame du signal vidéo numérique de dimension égale à celle de la matrice ;
- les moyens de répartition matricielle comportent une succession de conducteurs de signaux numériques en nombre égal à celui des lignes de la matrice de répartition, chacun avec des registres à décalage, connectés en série, imposant chacun un retard égal à l'écart temporel entre deux pixels successifs d'une ligne du signal vidéo numérique, la position d'un pixel réparti dans la matrice étant déterminé par un point de la matrice situé en amont d'un registre à retard, dont le nombre par ligne est égal au nombre de colonnes dans la matrice diminué d'une unité, et par un point en aval du registre à décalage le plus en aval ;

- lesdits moyens de repérage, dans ladite matrice, d'une zone en déplacement, par détection de la présence simultanée d'une valeur du signal binaire indiquant un dépassement et de la variation de la valeur de la constante de temps, comprennent des moyens pour déterminer, suivant des directions orientées discrètes numérisées, la pente de la variation de la valeur de la constante de temps au voisinage d'un pixel au centre de ladite matrice, constituant l'origine pour lesdites directions, et des moyens pour sélectionner la pente la plus forte de variation au voisinage de ce centre-origine et en déterminer la direction orientée, en tenant compte d'un critère de choix pour la sélection de la direction en cas de plus d'une direction de même pente maximale de variation, ces derniers moyens débitant, en sortie, des signaux représentatifs de la vitesse et de la direction orientée du déplacement dans la zone en déplacement, avec un signal de validation indiquant que ces signaux de vitesse et de direction sont validés, ainsi que de la valeur de la constante de temps.

De préférence dans le cas où le dispositif selon l'invention les comporte, les moyens formant les histogrammes comprennent chacun :

- des entrées recevant le signal dont on forme l'histogramme et un signal de validation en provenance des moyens de répartition matricielle ; et
- des moyens pour établir deux histogrammes unidimensionnels linéaires pour les deux coordonnées du plan et pour combiner ces deux histogrammes linéaires en un histogramme de surface représentant la zone de variation significative du signal d'entrée ; et
- une sortie débitant un signal représentatif de cette zone.

En outre les moyens formant les histogrammes comprennent de préférence :

- des moyens de calcul de changement de repère, dont les entrées reçoivent un signal de séquence de lignes, un signal de séquence de colonnes et un signal d'horloge de pixel et dont la sortie est représentative du changement de repère ;
- deux moyens de formation d'histogrammes pour deux axes, recevant les deux signaux de repère et formant les histogrammes de ces axes ; et
- des moyens de zone recevant les sorties des deux moyens de formation d'histogrammes pour les deux axes et débitant, en sortie, un signal d'information de pente globale des deux axes.

Dans certaines applications, on peut prévoir que ladite constante de temps est constituée par le numéro d'ordre des intervalles, en nombre limité, et en grandeur progressivement croissante, selon lesquels on découpe la valeur absolue de la différence entre la valeur de pixel actuel et la valeur juste antérieure du même pixel après lissage, pour chaque position de pixel.

Avantageusement pour déterminer le signal binaire de dépassement ou non-dépassement du seuil, on compare à un seuil, pour chaque position de pixel, la valeur absolue de la différence

entre la valeur du pixel actuel qui entre et la valeur du même pixel lissée juste antérieure en provenance de la mémoire.

En ce qui concerne le signal numérique d'amplitude, on le génère, de préférence, sous la forme d'un nombre entier qui représente la tendance au rapprochement entre la valeur du pixel actuel et la valeur de ce même pixel juste antérieure lissée, pour chaque position de pixel.

Une des caractéristiques de l'invention est constituée par le fait que, pour déterminer tant le signal binaire que le signal numérique d'amplitude, on utilise, pour la valeur du pixel juste antérieure, une valeur lissée de celui-ci mise en mémoire, afin de réduire les variations temporelles excessives de ce signal de pixel qui peuvent exister dans le signal vidéo numérique d'entrée en provenance de la caméra vidéo ou autre dispositif d'observation à sortie numérique.

On sait en effet qu'une opération de lissage a pour effet de substituer progressivement, à un signal numérique à variations importantes d'amplitude dans le temps, un signal présentant des variations plus réduites et donc plus facilement quantifiables avec un nombre réduit de paliers, donc de bits, dans le signal numérique d'amplitude lissé.

On va décrire maintenant, avec référence au dessin annexé, un mode de réalisation préféré, donné à titre d'exemple illustratif, et nullement limitatif, d'un dispositif selon l'invention mettant en oeuvre le procédé selon l'invention.

Sur ce dessin :

La Fig. 1 illustre très schématiquement l'ensemble du système selon l'invention avec son entrée et ses sorties, ainsi que le signal d'entrée pour ce système.

La Fig. 2 représente, sous forme de blocs fonctionnels, les principales unités d'un dispositif selon l'invention formant un ensemble de traitement temporel et spatial.

La Fig. 3 et la Fig. 4 illustrent les schémas fonctionnels, respectivement de l'ensemble de traitement temporel par calcul et de l'ensemble de traitement spatial par répartition matricielle, qui font partie du dispositif de la Fig. 2.

La Fig. 5 schématise le traitement temporel et le traitement spatial dans un système selon l'invention.

La Fig. 6 illustre la valeur numérique, selon le code de Freeman, des directions discrètes à partir d'un centre d'origine dans la matrice de la Fig. 4.

La Fig. 7 illustre deux matrices emboîtées de petite taille à l'intérieur de la matrice de répartition temporelle.

Les Fig. 8 et 9 décrivent deux autres types de matrice, respectivement hexagonale et en L renversé.

La Fig. 10 illustre schématiquement l'assemblage, selon la ligne $Z - Z^0$ du dispositif de la Fig. 2 avec un ensemble supplémentaire selon une réalisation préférée de l'invention.

La Fig. 11 représente, sous forme de blocs, ledit ensemble supplémentaire de la Fig. 10; la Fig. 2 et la Fig. 11 étant assemblées le long de la ligne $Z - Z^0$, en traits mixtes, figurant sur la Fig. 2 et la Fig. 10.

La Fig. 12 illustre la formation de deux histogrammes unidimensionnels et, à partir de ceux-ci, d'un histogramme superficiel de zone en déplacement, pour un signal d'entrée.

La Fig. 13 représente plus en détail un bloc ou unité de formation et traitement d'histogramme et son bloc de combinaison linéaire associé.

La Fig. 14 illustre un histogramme unidimensionnel.

La Fig. 15 et la Fig. 16 illustrent l'utilisation de la variation de la pente d'observation d'une scène.

La Fig. 17, la Fig. 18 et la Fig. 19 illustrent, sous forme de blocs pour les deux premières figures, d'autres applications possibles d'un dispositif selon l'invention, la première figure concernant la vidéoconférence, la seconde étant relative à la surveillance des autoroutes (ou routes principales) et la troisième concernant la commande d'une machine par le mouvement de la main d'un opérateur.

Les Fig. 20 et 21 représentent schématiquement l'application de l'invention à la surveillance de l'endormissement d'un conducteur automobile.

La Fig. 22 montre l'image transformée suivant le diagramme de Mallat.

En référant d'abord à la Fig. 1, on voit que, dans le mode de réalisation préféré, le dispositif 11, selon l'invention, comporte d'abord une entrée 12 qui reçoit un signal vidéo numérique S d'une caméra vidéo ou caméscope 13, à un ou plusieurs capteurs CCD de type CMOS, avec sortie numérique directe ou à sortie analogique convertie, dans un convertisseur analogique/numérique, en sortie numérique. Ce signal S est constitué, à la manière connue, par une succession de paires de trames entrelacées telles que TR_1 et TR'_1, TR_2 et TR'_2, comportant chacune une succession de lignes de balayage horizontal, chaque ligne (telle que $l_{1.1}$, $l_{1.2}$, ... $l_{1.17}$... de TR_1 et $l_{2.1}$ de TR_2) étant constituée par une succession de signaux élémentaires de pixels ou points-images PI représentatifs de points (tels que : $a_{1.1}$, $a_{1.2}$, et $a_{1.3}$ pour la ligne $l_{1.1}$; $a_{2.1}$, $a_{2.2}$ pour la ligne $l_{1.2}$; $a_{17.1}$ et $a_{17.2}$ pour la ligne $l_{1.17}$; $a_{1.1}$, $a_{1.2}$ pour la ligne $l_{2.1}$) de la scène 13a surveillée par la caméra 13 ; c'est pourquoi sur le dessin on a indiqué S(PI), à savoir un signal S constitué par des pixels PI.

A la manière connue, S(PI) comporte des signaux de synchronisation trame ST au début de chaque trame et de synchronisation ligne SL au début de chaque ligne.

On voit donc que le signal S(PI) est constitué

- par une succession de séquences (les trames successives) dans le cadre d'un domaine temporel et
- dans chaque séquence (dans chaque trame) par une série de sous-séquences (lignes, pixels) dans le cadre d'un domaine spatial.

Dans le domaine temporel, on désignera par l'expression «trames successives» les trames successives de même nature (c'est-à-dire les trames impaires, par exemple TR_1 , ou respectivement paires, par exemple TR'_1) des paires de trames, par exemple TR_1 - TR'_1, formant les images successives du signal vidéo numérique S(PI) et par l'expression «pixels successifs dans la même position» les valeurs de pixel (PI) successives en un même emplacement des trames successives de même nature, par exemple $a_{1,1}$ de $l_{1,1}$ de la trame TR_1 et $a_{1,1}$ de $l_{2,1}$ de la trame correspondante suivante TR_2.

Le dispositif 11 comporte par ailleurs des sorties 14 débitant divers signaux numériques, élaborés par lui, utiles pour signaler l'existence d'une zone ou «objet» (au sens le plus général indiqué ci-dessus) en déplacement relatif et sa localisation, de même que sa vitesse et sa direction de déplacement si celui-ci est effectif par rapport à un environnement sensiblement immobile, à savoir le signal complexe ZH groupant schématiquement les signaux signalant d'existence et la localisation de cette zone ou objet, la vitesse V et de direction orientée DI du déplacement, et éventuellement le signal vidéo numérique d'entrée S retardé en synchronisme avec les signaux précédents, pour tenir compte de leur temps de calcul, ce signal retardé SR permettant de visionner, sur l'écran d'un moniteur ou d'un téléviseur 10, l'image perçue par la camera 13 au moment où sont disponibles les renseignements concernant la zone éventuelle en déplacement relatif, à savoir le signal ZH, (V, DI), utilisables dans un ensemble de traitement et contrôle 10a.

Avec référence à la Fig. 2, on va expliciter la structure de la première partie du dispositif 11 de la Fig. 1, cette première partie étant représentée à l'intérieur du cadre 11a en traits interrompus de cette Fig. 2.

L'ensemble 11a comprend essentiellement, d'une part, une unité de traitement temporel 15, avec une unité mémoire associée 16, et, d'autre part, une unité de traitement spatial 17, avec une unité à retards 18 et une unité de séquencement 19 associées, ainsi qu'une horloge de pixel 20 cadençant l'unité de traitement temporel 15 et l'unité de séquencement 19.

L'unité 15 de traitement temporel, qui réalise, entre autres, un lissage du signal vidéo :

- élabore, à partir du signal vidéo numérique S, provenant de la caméra vidéo 13 et comportant une succession de valeurs de pixel PI, et à partir des impulsions HP engendrés par l'horloge 20 (à partir du signal S) à la cadence des pixels dans une trame (en particulier 13,5 MHz), un

certains nombre de grandeurs, comme exposé ci-après avec référence à la Fig. 3, et échange avec la mémoire 16 les valeurs de deux de ces grandeurs, à savoir les valeurs lissées L du signal vidéo numérique et les valeurs C de la constante de temps du lissage, les valeurs L et C étant suivies de la lettre O pour les valeurs entrant dans la mémoire 16 à partir de l'unité 15 ou de la lettre I pour les valeurs sortant de la mémoire 16 pour atteindre l'unité 15, et

- débite, en sortie, un signal binaire DP de dépassement ou non-dépassement de seuil et un signal numérique CO indiquant la valeur calculée actualisée de la constante de temps, à savoir la valeur CO envoyée dans la mémoire 16.

La structure par blocs de calcul et/ou comparaison de l'unité de traitement temporel 15 est explicitée sur la Fig. 3 ; l'unité 15 comporte quatre blocs 15a, 15b, 15c, 15d.

Le premier bloc 15a de l'unité 15, à partir

- du signal vidéo numérique d'entrée S, constitué formé par une succession de signaux de pixels PI, et
- d'une valeur lissée LI de ce signal S pour la trame correspondante juste antérieure, calculée auparavant par l'unité 15 en tant que LO et stockée temporairement dans la mémoire 16 (comme expliqué ci-après)
- à la cadence imposée par les signaux d'horloge HP de l'horloge 20, calcule la valeur absolue AB de la différence entre les valeurs entrantes de PI et LI pour une même position de pixel (par exemple $a_{1,1}$, de $l_{1,1}$ de TR_1 et de $l_{2,1}$ de TR_2)

$$AB = |PI - LI|.$$

Le deuxième bloc 15b est un bloc de test :

- il reçoit le signal numérique précité AB de l'unité 15a et un signal numérique de valeur de seuil SE, qui pourrait être fixe, mais qui est en général fonction de la valeur de pixel ; on le fait alors varier dans le même sens que celui-ci pour constituer une correction de gamma (les moyens connus afin d'effectuer la variation de SE pour réaliser une correction de gamma étant représentés par le bloc optionnel 15e en traits interrompus) ; et
- il compare ces deux signaux numériques représentatifs de AB et SE afin de déterminer un signal binaire DP, c'est-à-dire pouvant prendre deux valeurs 1 et 0, qui signalent un dépassement ou un non-dépassement, respectivement, dudit seuil SE par AB :
- si AB est supérieur à SE, DP se verra attribuer, dans l'unité 15b, la valeur 1 représentative d'un dépassement ;
- si AB est inférieur ou égal à SE, DP recevra, dans l'unité 15b, la valeur 0 représentative de non-dépassement.

En fait lorsque DP = 1, il y a une trop grande différence entre PI et LI, c'est-à-dire entre le signal vidéo numérique d'entrée et le signal vidéo numérique précédent lissé, et il faut diminuer cette différence en réduisant la constante de temps de lissage et inversement si DP = 0 il faut augmenter cette constante de temps.

Le troisième bloc 15c réalise justement la variation désirée de valeur de la constante de temps en fonction de la valeur de DP :

- si DP = 1, le bloc 15c diminue d'une valeur unitaire U la contante de temps : CO (nouvelle valeur de cette constante) = CI (ancienne valeur de la constante) - U ;
- si DP = 0, le bloc 15c augmente de la même valeur unitaire U la constante de temps : CO = CI + U.

A cet effet le bloc 15c reçoit, sur une entrée, le signal de dépassement binaire précité DP en provenance du bloc 15b et, sur une autre entrée, le signal CI, qui est la valeur de la constante de temps antérieure stockée dans la mémoire 16, et effectue la diminution ou l'accroissement, de la valeur unité U, de la constante de temps entrante CI qui devient CO envoyée dans ladite mémoire 16 en remplacement de CI.

Avantageusement la constante de temps, dont dépend la convergence du lissage (en fonction du temps nécessaire pour que la valeur lissée atteigne la valeur d'entrée du signal vidéo numérique), est représentée par une puissance de 2, à savoir par une valeur 2^p, et c'est alors ce nombre entier p qui sera diminué ou augmenté, dans le bloc 15c, d'une unité, c'est-à-dire de 1 ; alors sur la Fig. 3 on a U = 1 pour p,

- si DP = 1, le bloc 15c soustrait une unité (1) du facteur p de la constante de temps 2^p, qui devient 2^{p-1} ;
- si DP = 0, le bloc 15c ajoute une unité (1) au facteur p de la constante de temps 2^p, qui devient 2^{p+1}.

Le choix d'une constante de temps du type 2^p a le double avantage de correspondre à la physiologie de la vision humaine et de permettre des calculs plus simples, ce qui simplifie la structure du bloc 15c.

Le bloc 15c doit assurer en outre une double condition, à savoir garder CO entre deux valeurs limites : CO ne doit pas devenir négatif ($CO \geq 0$) et ne doit pas dépasser un seuil N ($CO \leq N$). Dans le cas particulier où CI et CO sont de la forme 2^p, le seuil supérieur N est représenté par un nombre entier n qui constitue une valeur maximale pour p.

Le seuil supérieur N (ou n) peut être soit constant, soit variable ; dans ce dernier cas une unité optionnelle 15f (en traits interrompus) réalise cette variation de N (ou n) sur l'ordre de l'utilisateur par exemple. Une augmentation de N a pour conséquence d'augmenter la sensibilité

de la détection du déplacement, tandis que la diminution de N améliore la détection des grandes vitesses.

Enfin le quatrième bloc 15d reçoit, sur une première entrée, la valeur CO de la nouvelle constante de temps élaborée dans le bloc 15c, sur une deuxième entrée, le signal vidéo numérique d'entrée S sous la forme d'une information de valeur de pixel PI et, sur une troisième entrée, la valeur lissée du signal vidéo numérique d'entrée précédent, à savoir LI, en provenance de la mémoire 16 et il calcule

$$LO = LI + (PI - LI) / CO$$

qui est débité sur sa sortie.

En fait le terme (PI - LI) / CO représente la modification apportée à la valeur lissée du signal vidéo numérique, en tenant compte de la valeur modifiée CO de la constante de temps et il est proportionnel à la différence algébrique entre la valeur effective du pixel d'entrée actuel PI en provenance de la caméra 13 et sa valeur de lissage antérieure LI et inversement proportionnel à CO.

Si $CO = 2^p$, alors

$$LO = LI + (PI - LI) / 2^{po}$$

en tenant compte de *po*, valeur de *p* calculée dans l'unité 15c, et qui remplace dans la mémoire 16 la valeur précédente *pi* de *p*.

Par conséquent l'unité de traitement temporel 15 à quatre blocs de calcul 15a, 15b, 15c, 15d :

- reçoit S(PI) de la caméra vidéo 13, les impulsions d'horloge HP, pour le cadencement des opérations, les signaux de seuil SE et N (ou *n*) ;
- détermine, à partir de signaux d'entrée LI et CI en provenance de la mémoire associée 16, des signaux actualisés LO et CO qui sont envoyés dans ladite mémoire en remplacement respectivement de LI et CI et qui représentent respectivement les nouvelles valeurs calculées du signal vidéo numérique lissé et de la constante de temps ;
- et débite en sortie, vers l'unité de traitement spatial 17, à travers l'unité à retards 18, le signal CO précité et le signal binaire de dépassement DP qu'elle a calculés à partir de PI, LI et SE.

L'opération de lissage a pour but de normaliser les variations de la valeur numérique du signal vidéo d'entrée pour chaque pixel ou point d'image, à savoir la variation de chaque PI, en réduisant les écarts de variation et en substituant, pour chaque pixel, aux valeurs successives réelles variables de PI en ce point d'image des valeurs lissées LO, moins variables que les valeurs PI.

Ainsi à chaque PI entrant, l'unité de traitement temporel 15 substitue, en combinaison avec la mémoire 16, une valeur lissée LO à variations réduites, par mise en oeuvre d'un signal binaire DP de dépassement ou non d'un seuil et d'un signal CO de constante de temps qui sont actualisés et envoyés dans l'unité de traitement spatial 17 illustrée sur la Fig. 4.

On peut repérer chaque pixel à l'intérieur de la surface d'une trame par deux coordonnées (en principe orthogonales) d'abscisses et d'ordonnées, à savoir x et y, en attribuant au pixel deux indices i (numéro de la ligne) pour la coordonnée y et j (numéro du pixel dans la ligne) pour la coordonnée x. Chaque pixel d'indices i et j a une valeur vidéo (valeur de l'amplitude du signal vidéo) PI_{ij}.

Si on considère maintenant l'évolution, dans le temps t, des PI_{ij} pour des trames correspondantes successives aux instants successifs t_0, t_1, t_2, t_3 ... , séparés par une période T correspondant à la période d'image (égale en général à deux périodes de trame) et qui peut être de 0,04 s dans le cas d'une fréquence de 25 Hz pour le signal d'image vidéo et de 0,0333...s dans le cas d'une fréquence de 30 Hz pour ce signal, ou de 50 Hz pour des capteurs à image progressive (1 trame/1 image), le signal d'un pixel vidéo d'indices de localisation i et j a des valeurs successives notées PI_{ijt0}, PI_{ijt1}, PI_{ijt2}, PI_{ijt3} ... à ces instants t_0, t_1, t_2, t_3 ...

Dans le cadre de l'invention on substitue, dans l'unité de traitement temporel 15, à P_{ijt} les valeurs de lissage successives LO_{ijt}, savoir LO_{ijt0}, LO_{ijt1}, LO_{ijt2}, LO_{ijt3} ...

Pour chacun des pixels ou points d'image successifs de coordonnées i, j, à l'instant t, c'est-à-dire P_{ijt}, on substitue à sa valeur réelle PI_{ijt} une valeur lissée donnée par la formule

$$LO_{ijt} = LI_{ijt0 \,..\, t1} + (PI_{ijt} - LI_{ij\,t-\,t1}) / CO_{ijt}$$

La constante de temps est de préférence de la forme 2^{pijt}.

Les calculs effectués dans l'unité 15, notamment dans le bloc 15d, pour chaque intervalle de temps T séparant t_1 de t_0, t_2 de t_1, t_3 de t_2, etc. assurent une convergence de la valeur LO_{ijt} vers la valeur PI_{ijt}, dont la rapidité dépend de la constante de temps qui est variable dans l'espace (et dépend donc de i et j) et dans le temps (et dépend donc de t) et qu'on peut écrire CO_{ijt}.

Dans tous les cas plus CO_{ijt} est grand, plus la convergence de LO_{ijt} est lente. Si $CO_{ijt} = 1$, il n'y a plus de lissage.

On peut évidemment dans les formules précédentes substituer respectivement les coordonnées cartésiennes y et x aux indices i et j de numéro d'ordre de ligne et de pixel par ligne.

Le bloc 15a calcule $AB = |PI - LI|$, les indices i, j et t n'étant pas explicités, la grandeur AB étant représentative de la variabilité instantanée du signal vidéo numérique S(PI) par rapport au signal lissé LI, pour chaque triplet i, j, t.

C'est une des caractéristiques d'un dispositif selon l'invention de réaliser, dans l'unité 15 de traitement temporel, un lissage des valeurs de pixel, de déterminer les valeurs successives, pour chaque pixel, d'une constante de temps de lissage et d'un signal binaire de dépassement ou non-dépassement d'un seuil par la valeur absolue de la différence entre la valeur du pixel et la valeur lissée de ce pixel pour deux trames correspondantes successives, pour distribuer, suivant une matrice plane, à la fois les valeurs numériques, à un même instant, de ladite constante de temps et les valeurs dudit signal binaire pour les pixels d'une portion limitée de trame, portion qui balaie la trame, afin de localiser, grâce aux variations locales de ces deux valeurs à cet instant, une zone en mouvement relatif et de déterminer la vitesse et la direction (orientée) de déplacement effectif en fonction de cette répartition, ceci étant effectué un ensemble de traitement spatial, décrit ci-après avec référence à la Fig. 4.

La convergence du lissage est assurée par les blocs de calcul 15b et 15c qui déterminent une variation de la valeur de la nouvelle constante de temps entrante CO (en fait CO_{ij}) telle qu'elle accélère la convergence. Ceci est réalisé en comparant AB (en fait AB_{ij}) à un seuil SE, qui peut avoir une valeur constante ou de préférence variable et dans ce dernier cas être fonction de la valeur de pixel afin d'assurer une correction de gamma : si la valeur de pixel croît, le seuil croît et vice et versa. Le résultat de la comparaison de AB avec le seuil SE (en fait SE_{ij} si le seuil varie en fonction de la valeur de pixel, cette variation étant réalisé dans un bloc éventuel de calcul 15e) génère dans le bloc 15b le signal binaire de déplacement DP à deux valeurs possibles 1 et 0, comme expliqué ci-dessus.

Le signal binaire DP arrive dans le bloc 15c pour faire varier la valeur de la constante de temps. A cette fin l'unité de calcul 15c reçoit également la valeur de la constante de temps entrante CI en provenance de la mémoire 16 et l'actualise en nouvelle valeur de la constante de temps CO qui est, entre autres, envoyée dans la mémoire 16 où elle est substituée à l'ancienne valeur CI ; en fait pour les valeurs de CI et CO, il s'agit de CI_{ij} et CO_{ij} à deux instants successifs, par exemple t_0 et t_1, séparés par l'intervalle T entre deux trames correspondantes (soit impaires, soit paires) successives.

Le bloc 15c recevant DP et CI ajoute ou soustrait une valeur unitaire (U) à la valeur de la constante de temps CI ou une unité à p lorsque CO est de la forme 2^p suivant que le signal binaire DP représente un dépassement (DP = 1) ou un non-dépassement (DP = 0).

S'il y a dépassement, c'est que la valeur de cette constante est trop forte et on le diminue et vice et versa.

En outre le bloc 15c s'assure que la nouvelle valeur de la constante de temps CO, déduite de CI par addition ou soustraction d'une unité, reste comprise entre 0 (CO non négatif) et une

valeur seuil N (CO non supérieur à N). Si cette double condition n'est pas remplie, le bloc 15c ne modifie pas la valeur de CI (qui était elle effectivement à l'intérieur de l'intervalle de 0 à N, limites comprises) et alors CO = CI.

La valeur limite supérieure N ou *n* peut ou bien être constante ou bien être variable ; dans ce dernier cas elle ne doit pas dépasser une valeur limite N max (ou *n* max) ; la variation éventuelle étant imposée par le bloc 15f sous l'effet d'une commande à la disposition de l'utilisateur.

En variante, on peut faire dépendre N ou *n* de PI (N, *n* et PI étant en fait affectés du triplet d'indices *i, j, t*) afin d'assurer une régulation de la variation de LO (qui est calculé dans le bloc 15d) en fonction du niveau de PI, ce qui peut s'exprimer par «N_{ijt} ou n_{ijt} est une fonction de PI_{ijt}», la détermination de N_{ijt} ou $n_{ijt} = f(PI_{ijt})$ étant effectuée dans un bloc de calcul substitué au bloc illustré 15f et recevant, en plus de N, la valeur de PI à partir de la caméra vidéo 13.

On peut avantageusement imposer à la constante de temps C (en fait à chaque C_{ijt}) la condition d'avoir une valeur numérique qui soit un multiple de 2, plus particulièrement une puissance de 2, quel que soient *i, j* et *t* ; dans ce cas particulier $C_{ijt} = 2^{p(ijt)}$, *p* étant un nombre entier petit, fonction de *i, j* et *t*, représentable par un nombre réduits de bits. Cette condition procure les avantages supplémentaires mentionnés ci-dessus :

- la loi de convergence du lissage est voisine des lois de la physiologie de la vision humaine
- la réalisation électronique des blocs 15c et 15d est plus simple ; en particulier, dans le bloc 15d, qui a pour objet de déterminer l'évolution de la valeur de la constante de temps, pour chaque couple *i, j*, par la formule

$$LO = LI + (PI - LI) / CO,$$

les calculs sont simplifiés si CO est de la forme 2^p (*p* étant un nombre entier petit), la valeur de seuil *n* de *p* étant elle-même un nombre entier petit, représentable par un nombre limité de bits.

Dans tous les cas, la nouvelle valeur LO de lissage du signal vidéo numérique d'entrée S est envoyée dans la mémoire 16 dans laquelle elle se substitue à LI (et ceci pour chaque couple d'indices *ij*).

On peut constater, à l'examen des Fig. 2 et 3 que l'unité de traitement temporel 15, qui comporte les blocs de calcul 15a, 15b, 15c, 15d, et éventuellement le bloc 15e et/ou le bloc 15f, et qui coopère avec la mémoire 16, détermine, comme exposé ci-dessus, et délivre en sortie, pour chaque triplet *i, j, t*,

- d'une part, la valeur lissée LO actualisée qui est transférée à la mémoire 16 en remplacement de la valeur de lissage antérieur LI

– d'autre part, deux signaux numériques, à savoir :

– un signal binaire DP, qui indique soit le dépassement (DP = 1), soit le non-dépassement (DP = 0) d'un certain seuil par la variation en valeur absolue du signal de pixel entrant en provenance de la caméra vidéo par rapport au signal de pixel antérieur lissé pour le même point ; et

– un signal numérique d'amplitude, constitué par la valeur de la constante de temps actualisée CO ;

ces signaux numériques DP et CO sont reçus, à travers l'unité à retards 18, par l'unité de traitement spatial 17 qui sera décrite ci-après, le signal CO étant également reçu par mémoire 16 dans laquelle la valeur CO se substitue à la valeur antérieure CI pour le même pixel.

On voit donc que la capacité de la mémoire 16 pour stocker les valeurs successives du signal de pixel lissé, d'une part, et de la constante de temps, d'autre part, c'est-à-dire, en supposant qu'il y ait R pixels dans une trame, donc 2R pixels par image complète, doit être d'au moins $2R(e+f)$ bits en appelant e et f les nombres de bits attribués respectivement à un signal de pixel et à une constante de temps. En fait la capacité de la mémoire n'a pas besoin d'être beaucoup plus grande ; il suffit qu'elle dépasse $2R(e+f)$ bits du nombre de bits nécessaire pour assurer son fonctionnement correct, notamment pour l'adressage et l'extraction des bits des signaux de pixel lissés et des bits des constantes de temps en fonction des indices i et j. Si chaque image vidéo est constituée par une seule trame il suffira de $R(e+f)$ bits au lieu de $2R(e+f)$ bits.

Les sorties DP_{ij} et CO_{ij}, à un instant t, de l'ensemble de traitement temporel 15 sont analysées et utilisées dans un ensemble de traitement spatial illustré sur la Fig. 4, l'assemblage des Fig. 3 et 4 étant illustré sur la Fig. 2.

En fait l'ensemble 15 de traitement temporel traite les signaux de trames, tandis que l'ensemble 17 de traitement spatial de la Fig. 4 traite les sous-séquences de lignes et de pixels dans une trame.

Sur la Fig. 5 on a schématisé le traitement temporel des séquences de trames correspondantes successives TR_1, TR_2, TR_3 superposées sur la figure et le traitement spatial dans la première de ces trames, c'est-à-dire TR_1, en illustrant les coordonnées cartésiennes x et y et un pixel PI de coordonnées y, x, c'est-à-dire d'indices i, j au temps t_1 ; les pixels successifs de mêmes indices ij sur les trois trames TR_1, TR_2, TR_3 sont indexés i,j,t_1, ijt_2 et ijt_3 respectivement et ils ont les valeurs de pixels $PI_{ij,1}$, PI_{ijt2} et PI_{ijt3} respectivement. Un plan de la Fig. 5 correspond à un traitement spatial d'une trame, tandis que la superposition des plans correspond au traitement temporel (le temps t étant la variable)

L'unité de traitement spatial 17, à laquelle est associé une unité à retards 18 (également illustrée sur la Fig. 4), coopère avec une unité de commande 19 qui est contrôlée par une horloge 20 qui débite une impulsion d'horloge HP à chacun des signaux de pixel successifs (Fig. 2 pour l'ensemble)

Les sorties DP_{ij} et CO_{ij} de l'unité 15 de traitement temporel sont réparties dans l'unité 17 suivant une matrice 21 de dimension réduite comportant un nombre de lignes et un nombre de colonnes très inférieurs respectivement au nombre de lignes L et de pixels M par ligne des DP_{ij} et CO_{ij} à un instant t donné. En particulier la matrice peut comprendre $2l + 1$ lignes suivant l'axe des y et $2m + 1$ colonnes suivant l'axe des x (en coordonnées cartésiennes), l *et* m étant des nombres entiers petits. Avantageusement on choisit l et m parmi les puissances de 2, l étant égal à 2^a et m à 2^b, a et b étant des nombres entiers de l'ordre de 2 à 5 par exemple. Pour simplifier le dessin et l'explication, on prend, à titre d'exemple, $m = l$ (bien qu'ils puissent être différents) et $m = l = 2^3 = 8$; dans ce cas la matrice 21 aura $2 \times 8 + 1 = 17$ lignes et 17 colonnes.

Sur la Fig. 4 on a représenté une partie des 17 lignes Y_0, Y_1, ... Y_{15}. Y_{16} et une partie des 17 colonnes X_0, X_1, ... X_{15}.X_{16} constituant la matrice 21 de l'unité 17.

Il s'agit de répartir suivant la matrice 21 à $l + 1$ lignes, en particulier 17 lignes, et $m + 1$ colonnes, en particulier 17 colonnes, les flux entrants des DP_{ij} et CO_{ij}, c'est-à-dire des signaux binaires de dépassement DP et des signaux numériques d'amplitude représentant la constante de temps CO, qui arrivent de l'unité 15 de traitement temporel, suivant une distribution matricielle plus étendue pour une trame, à savoir de L lignes, en particulier 312,5, et M pixels par ligne, en particulier de l'ordre de 250 à 800 suivant le standard TV utilisé.

Pour distinguer les deux matrices à savoir celle du signal vidéo de L x M et celle de l'unité 17 de l x m, référencée 21, on utilisera les indices i et j suivant les deux coordonnées de la première (qui n'apparaît que lorsqu'on visualise sur un écran de téléviseur ou moniteur le signal vidéo numérique) et les indices x *et* y suivant les deux coordonnées de la seconde (représentée sur la Fig. 4) ; à un instant donné, un pixel de valeur instantanée PI_{ij} est caractérisé à l'entrée de l'unité de traitement spatial 17 par deux signaux numériques DP_{ij} et CO_{ij}. La matrice du L x M de ces deux signaux se déplace par balayage à travers la matrice 21 de $(2l + 1) \times (2m + 1)$ beaucoup plus petite, comme expliqué ci-après avec référence à la Fig. 4, la matrice 21 matérialisant $(2l + 1) \times (2m + 1)$ pixels correspondant à une même trame.

Dans cette matrice 21, chaque pixel est défini par un numéro d'ordre de ligne compris entre 0 et 16 (bornes incluses) pour les lignes Y_0 à Y_{16} respectivement et un numéro d'ordre de colonne compris entre 0 et 16 (bornes comprises) pour les colonnes X_0 à X_{16} respectivement, dans le cas où $l = m = 8$, soit $2l + 1 = 2m + 1 = 17$. Dans ce cas la matrice 21 réalisera une

représentation dans le plan de 17 x 17 = 289 pixels, alors que la matrice du signal vidéo comportera plusieurs dizaines ou centaines de milliers des pixels ou même d'avantage.

Sur la Fig. 4, on a illustré par des rectangles allongés horizontaux Y_0 à Y_{16}, dont seulement quatre ont été représentés, à savoir Y_0, Y_1, Y_{15} et Y_{16}, et par des lignes verticales X_0 à X_{16}, dont seulement quatre ont été représentées, à savoir X_0, X_1, X_{15} et X_{16}, cette matrice 21 (de l'unité 17) à 17 x 17 points d'image ou pixels d'indices définis à l'intersection d'une ligne d'ordonnée et d'une colonne d'abscisse. Par exemple la position de pixel P_{88} est à l'intersection de la colonne 8 et de la ligne 8, comme illustré sur la figure en g, centre de la matrice 21.

Pour réaliser la répartition spatiale successive de portions de cette matrice de L x M suivant la matrice 21 de $(2l + 1)(2m + 1)$, on associe à l'unité 17 une unité à retards 18 qui reçoit, d'une part, les signaux DP et CO (affectés des indices *ij*) et, d'autre part, le signal de pixel d'entrée S, c'est-à-dire PI (également d'indices *ij*), ainsi qu'un signal HP en provenance d'une horloge 20 et des signaux de séquence ligne SL et de séquence colonne SC (Fig. 2 et 4).

Comme représenté sur la Fig. 1, le signal S(PI) comporte, en plus des signaux de valeurs de pixel tels que $a_{i,1}$, $a_{i,2}$ constituant une séquence temporelle (trames successives) et des sous-séquences spatiales (pixels par lignes dans chaque trame), des signaux de synchronisation ST, SL, dont l'unité d'horloge 20 déduit non seulement un signal d'horloge, à la fréquence de 13,5 MHz par exemple, à savoir un top pour chaque pixel d'une trame vidéo, mais également des signaux de blanking BL qui rendent non-opératoire l'unité 19 pendant les signaux de synchronisation précités.

En réponse à ces signaux HP et BL en provenance de l'horloge 20 (Fig. 2) l'unité de cadencement 19 débite sur l'unité à retards 18 un signal de séquence lignes SL à une fréquence égale au quotient de 13,5 MHz par le nombre de colonnes par trame, soit par exemple 400, et un signal de trame SC dont la fréquence est égale au quotient précité 13,5/400 MHz divisé par le nombre de lignes d'image vidéo, par exemple 312,5, ainsi que HP.

C'est à partir de ces signaux SL et SC et du signal d'horloge HP que l'unité 18 assure la répartition spatiale ligne par ligne suivant la matrice 21.

A cet effet les lignes successives Y_0 à Y_{16} reçoivent les signaux DP et CO :

- non retardés (ligne Y_0) ;
- retardés d'une période TP, égale à la durée d'une ligne de trame (ligne Y_1),
- retardés de 2TP (ligne Y_2),

et ainsi de suite jusqu'à

- retardés de 16TP (ligne Y_{16}).

Les retards successifs de la durée d'une ligne de trame, savoir TP, sont réalisés dans une cascade de seize circuits à retard r_1, r_2, ... r_{16} qui desservent les lignes Y_1, Y_2 ... Y_{16} respectivement, la ligne Y_0 étant desservie directement par les signaux DP et CO non retardés en provenance de l'unité 15.

L'ensemble des circuits r_1, r_2, ... r_{16} peut être constitué par une ligne à retard à seize sorties, le retard imposé par une section quelconque entre deux sorties successives étant constant et égal à TP.

Le défilement de la totalité de la matrice trame de L x M sur la matrice 21 de $(2l + 1)(2m + 1)$ est assurée, en ce qui concerne les lignes successives des trames successives par roulement, par l'unité de cadencement 19 grâce aux signaux SL de séquence lignes, comme suit.

En ce qui concerne le dépiacement par roulement des pixels d'une ligne de la matrice d'une trame sur la matrice de 17 x 17, par exemple de X_0 à X_{16} sur la ligne Y_0, elle est réalisée par une cascade de seize registres à décalage d sur chacune des 17 lignes de Y_0 à Y_{16} (soit au total 16 x 17 = 272 registres à décalage) disposés dans chaque ligne entre deux positions successives de pixels, soit entre les positions PI_{00} et PI_{01} le registre d_{01}, entre les positions PI_{01} et PI_{02} le registre d_{02}, etc. Chaque registre impose un retard de TS égal à l'écart temporel entre deux pixels successifs d'une ligne, grâce aux signaux SC de séquence colonne.

On notera que du fait que les lignes l_1, l_2 ,...l_{17} d'une trame TR_1 (Fig. 1), tant pour S(PI) que pour DP et CO, arrivent décalées de TP (durée complète d'une ligne) l'une après l'autre à l'unité 18 et que celle-ci les distribue avec des retards croissant progressivement de TP sur les lignes Y_0, Y_1 ... Y_{17}, celles-ci affichent à un instant donné les signaux de DP et CO pour les lignes l_1 , l_2 , ... l_{17} d'une même portion de trame.

De même dans une ligne donnée, telle que l_1, les signaux de pixels successifs $a_{1,1}$, $a_{1,2}$... arrivent décalés de TS et les registres à décalage d imposent un retard également égal à TS ; il en résulte donc que les signaux de DP et CO pour des pixels d'une ligne donnée Y_0 à Y_{16} de la matrice 21, pixels disponibles sur cette ligne, sont contemporains, c'est à dire qu'ils correspondent à une même portion de trame.

Il s'agit donc, tant pour les lignes que pour les pixels de ces lignes d'une portion de trame, d'un traitement purement spatial, car la matrice 21 affiche, en ses 17 x 17 positions de pixels, les valeurs de DP et CO pour les 17 pixels de chacune des 17 lignes d'une même matrice du signal vidéo numérique S(PI), bien que ces pixels, tels que $a_{1,1}$, arrivent successivement ligne par ligne et pixel par pixel dans chacune des lignes successives (Fig. 1) dans l'unité 18, de même que les signaux correspondants DP et CO qui sont affichés.

Les signaux représentatifs des CO et des DP contemporains de la matrice 21 sont disponibles, à un instant donné, sur les $16 \times 17 = 272$ sorties des registres à décalage, ainsi qu'en amont des 17 registres en tête des 17 lignes, c'est-à-dire des registres $d_{0,1}$, $d_{1,1}$ ... $d_{16,1}$, ce qui fait au total $16 \times 17 + 17 = 17 \times 17$ sorties pour les 17×17 positions $P_{0,0}$, $P_{0,1}$, ... $P_{8,8}$ ... $P_{16,16}$.

A l'intérieur de la matrice 21, autour du centre de celle-ci g de coordonnées $x = 8$, $y = 8$ (et c'est pour cela que le nombre de lignes et le nombre de colonnes de la matrice 21 est de préférence impair : $2l + 1$ et $2m + 1$ respectivement), on peut considérer en particulier une petite matrice de 3 lignes et de 3 colonnes dont l'élément central de ses 9 éléments est justement le pixel g de coordonnées $x = 8$, $y = 8$. Soit

$$\begin{matrix} a & b & c \\ d & g & f \\ g & h & i \end{matrix} \quad \text{(M3)}$$

cette petite matrice, dont l'élément central g a été souligné.

A cette matrice de 3×3 éléments, comportant 8 emplacements a, b, c, d, f, g, h, i tout autour de l'élément ou emplacement central g, on peut faire correspondre 8 directions orientées partant chacune de l'emplacement central g et aboutissant à l'un des 8 autres.

A cet effet, on peut repérer les 8 directions au moyen du code de Freeman illustré sur la Fig. 6, les directions étant codées de 0 à 7, à partir de l'axe des x, de 45° en 45°. En code de Freeman, les 8 directions orientées possibles numérotées de 0 à 7, sont représentables par un nombre à 3 bits, car $2^3 = 8$, à savoir 8 possibilités.

Si on reprend la petite matrice précédente M3, les 8 directions selon le code de Freeman à partir de la position centrale g sont les suivantes

$$\begin{matrix} 3 & 2 & 1 \\ 4 & g & 0 \\ 5 & 6 & 7 \end{matrix}$$

comme explicité sur la Fig. 6.

Revenant à la matrice 21 de la Fig. 4 à 17×17 points d'image ou pixels, on va exposer d'abord comment on y repère une zone en déplacement relatif par rapport à un environnement sensiblement immobile dans la scène, observée par la caméra vidéo 13 et donc représentée dans le signal numérique vidéo S composé de pixels PI_{ij}, et comment on détermine la vitesse et la direction orientée d'un déplacement effectif par rapport à un environnement sensiblement immobile.

Entre deux trames successives, telles que TR_1 et TR_2 (Fig. 5), les PI_{ij} pixels du signal S seront caractérisés, en ce qui concerne leur variation entre l'instant t_1 (première trame) et l'instant t_2 (deuxième trame) par les deux signaux DP_{ij} et CO_{ij} repartis par balayage suivant la matrice 21.

Il existe une variation significative de la valeur de pixel en un point de cette matrice si DP = 1 pour ce point. Donc une zone en déplacement effectif est repérée par la zone de la matrice à l'intérieur de laquelle DP = 1 en chaque point.

En fait on examine simultanément, dans une unité de calcul 17a, à l'intérieur de la matrice 21 à 17 x 17 emplacements, les différentes matrices carrées emboîtées centrées sur *g* de dimensions 15 x 15, 13 x 13, 11 x 11, 9 x 9, 7 x 7, 5 x 5 et 3 x 3, cette dernière étant la matrice M3 précitée. On détermine quelle est la matrice parmi celles-ci qui est celle de plus petite dimension à posséder des DP = 1 alignés suivant une droite qui détermine la direction du déplacement de la zone dans laquelle DP = 1 et qui détermine une variation de + 1 et - 1 autour de CO. Pour une variation de + 1 et - 1 autour de CO, il faut que DP = 1 à chaque valeur, pour accepter le test. C'est la matrice la plus petite participant au test qui est choisie (ligne de plus grande pente).

Ensuite à l'intérieur de cette zone en déplacement dans une des matrices emboîtées, par exemple dans la petite matrice M3 de 3 x 3 éléments, on détermine si CO varie de chaque coté de la position centrale, dans une direction donnée, de + 1 dans une direction orientée et de - 1 dans la même direction mais de sens opposé. Par exemple si on a - 1, 0, + 1 dans la direction (orientée) 1, c'est-à-dire dans les positions *g, e, c* respectivement de la petite matrice M3, alors le déplacement existe dans cette matrice de droite à gauche dans la direction (orientée) 1 du code de Freeman (Fig. 6). Bien entendu simultanément dans cette direction de la petite matrice DP = 1. Plus CO varie de + 1 ou - 1 entre deux positions voisines suivant une direction dans une matrice de plus grande taille parmi les matrices emboîtées de 3 x 3 à 15 x 15, plus la vitesse du déplacement est grande. Par exemple si on a - 1, 0, + 1 dans la direction orientée 1, c'est-à-dire *g, e, c*, dans la matrice de 9 x 9, référencée M_9 le déplacement sera plus rapide que dans le cas où l'on a - 1, 0, + 1 dans la matrice M_3 de 3 x 3 (Fig. 7).

Du fait que CO est une puissance de 2 et est représentée par cette puissance dans les modes de réalisation préférés, on peut repérer une gamme étendue de vitesses en utilisant quelques bits seulement pour la puissance de 2, tout en repérant même une vitesse relativement faible (qu'on peut choisir en augmentant l'écart pour le test ; par exemple - 2, 0, + 2 dans la matrice M3 de 3 x 3 indique une vitesse deux fois plus faible par rapport à la vitesse correspondante à la matrice M3 de - 1, 0, + 1, et ce pour les mêmes positions *g ,e ,c*).

Il y a lieu de prévoir en outre deux tests pour lever les incertitudes :

- un premier test choisit la variation la plus forte, c'est-à-dire la constante de temps la plus élevée, dans le cas où l'on a dans une des matrices emboîtées, par exemple dans la petite matrice M3 de 3 x 3 éléments, des variations de CO suivant plusieurs directions ;
- un second test choisit, arbitrairement, une direction parmi deux (ou plusieurs) directions suivant lesquelles la variation de CO est identique, par exemple en choisissant la valeur la plus faible du code de Freeman ; en fait un tel cas se produit la plupart du temps lorsque la direction (orientée) réelle du déplacement se trouve sensiblement entre deux directions codées successives dans le code de Freeman, par exemple entre les directions 1 et 2, ce qui correspond à une direction (orientée) pouvant être notée 1, 5 (Fig. 6), faisant environ 67,5° avec celle de l'axe des *x* (direction 0 dans le code de Freeman).

La détermination de la direction orientée et de la vitesse d'une zone en déplacement effectif de la manière qui vient d'être décrite est réalisée par l'unité de calcul 17a (Fig. 4) associée à l'unité 17 et qui reçoit les 17 x 17 sorties précitées de la matrice 21, tant pour CO que pour DP (deux sorties de 21, à savoir $s_{8,1}$ et $s_{8,16}$, sont représentées). L'unité 17a traite les valeurs de CO et DP pour les matrices emboîtées successives et détermine d'après ces valeurs la direction orientée (dans le code de Freeman) et la vitesse du déplacement (suivant la matrice à retenir parmi les matrices emboîtées), en appliquant éventuellement les tests précités.

Le défilement de la totalité d'une trame du signal vidéo numérique à travers la matrice 21 a lieu

- d'abord pour le groupe des 17 premières lignes, lignes 1 à 17, de la trame : de la gauche vers la droite de la trame en considérant le mouvement relatif, comme illustré pour la trame TR_2 sur la Fig. 5 : de la portion TM_1 à l'extrême gauche, puis TM_2 décalée d'une colonne par rapport à TM_1 et ceci jusqu'à TM_{N1} (M étant le nombre de pixels par ligne de trame) à l'extrême droite ;
- ensuite, de manière analogue de la gauche vers la droite de colonne en colonne, pour les lignes 2 à 18 de la trame ; et
- ainsi de suite en descendant chaque fois d'une ligne : ligne 3 à 19... jusqu'au dernier groupe en bas de la trame, à savoir celui des lignes L - 16 ... L (L étant le nombre de lignes par trame).

En considérant les Fig. 2 et 4 on voit que les sorties des unités 17, 18 et 19, c'est-à-dire de l'ensemble de traitements spatial, sont :

- un signal V représentant la vitesse du déplacement, d'après l'amplitude de la variation maximale de CO dans la zone repérée, dont la valeur peut être par exemple représentée par une échelle de huit nombres entiers de 0 à 7 si la vitesse est sous la forme numérique de puissances de 2, donc comporte 3 bits ;

- un signal DI représentant la direction de ce déplacement, d'après la direction de cette variation maximale, la valeur de DI pouvant également être représentée par une parmi les huit valeurs de 0 à 7 dans le code de Freeman, donc comporter 3 bits ;
- un signal de validation VL précisant que le résultat pour la vitesse et la direction (orientée) est valide, afin de pouvoir distinguer une sortie valide avec V = 0 et DI = 0, de l'absence de sortie due à un incident, ce signal étant soit 1 (sortie valide) soit O (sortie absente) ; donc un seul bit est nécessaire pour VL ;
- avantageusement un signal CO de constante de temps, donc de 3 bits par exemple ;
- (les 3 ou 4 signaux V, DI, VL et éventuellement CO étant débités par l'unité 17 de traitement spatial et son électronique associée)
- un signal vidéo retardé SR constitué par le signal vidéo d'entrée S retardé dans l'unité à retards 18 des 16 durées consécutives de lignes TR et donc de la durée de la répartition du signal S dans la matrice 21 de 17 x 17, de manière à disposer du signal vidéo numérique contemporain de la représentation matricielle dans la matrice 21, signal dont le contenu peut être affiché en clair sur l'écran d'un téléviseur ou d'un moniteur ;
- l'ensemble des trois signaux de sortie de l'unité 19, à savoir les signaux d'horloge HP, de séquence ligne SL et de séquence de colonne SC.

Aux matrices rectangulaires emboîtées des Fig. 4 et 7, on peut substituer des matrices hexagonales emboîtées (Fig. 8) ou une matrice en forme de L renversé (Fig. 9).

Dans le cas de la Fig. 8, les matrices emboîtées (dont on n'a représenté que les matrices MR1 et MR2 les plus centrales) sont toutes centrées sur le point MR0 qui correspond au point central (en lequel le signal binaire est «0») des matrices M3, M9 de la Fig. 7. L'avantage d'un système de matrices hexagonales c'est qu'il permet, d'une part, d'utiliser des axes de coordonnées obliques, x_n, y_n et, d'autre part, une décomposition en triangles ayant des cotés identiques, ce qui réalise le calcul des vitesses isotropes.

La matrice de la Fig. 9 est composée d'une seule ligne L_n et d'une seule colonne C_n à partir de la case centrale MR_n dans laquelle les deux signaux DP et CO valent respectivement «1» pour DP et augmente ou diminue d'une unité pour CO, en cas de déplacement.

On détermine ainsi si la direction du déplacement (relatif) est

- dans le sens de la coordonnée x : le signal CO est identique dans toutes les cases de la colonne C_n,et le signal binaire DP est égal à 1 dans les cases de la ligne L_n , de l'origine MR_n à la valeur CO_n jusqu'à la case où CO vaut $CO_n + 1$ ou $- 1$ inclus ;

- dans le sens de la coordonnée y : le signal CO est identique dans toutes les cases de la ligne L_n, et le signal binaire DP est égal à 1 dans les cases de la colonne C_n, de l'origine MR_n à la valeur CO_n jusqu'à la case où CO vaut $CO_n + 1$ ou $- 1$ inclus ;
- ou enfin oblique relativement à x et y : le signal binaire DP est égal à 1 et CO vaut CO_n dans les cases de L_n et dans les cases de C_n, la pente étant déterminée par la perpendiculaire à la ligne passant par les deux cases dans lesquelles le signal CO_n change de valeur d'une unité, le signal DP étant toujours égal à 1.

Sur la Fig. 9 on a illustré le cas où DP = 1 et CO_n change de valeur d'une unité dans deux cases particulières L_{n5} et C_{n5} et indiqué la pente correspondante P_p.

Dans tous les cas, la vitesse du déplacement est fonction de la case dans laquelle CO change de valeur d'une unité.

Si CO change d'une unité que dans L_n ou C_n, elle correspond à la valeur de la case de variation de CO.

Si CO change d'une unité dans une case de L_n et dans une case de C_n, la vitesse est proportionnelle à la distance entre MR_n et E_x (intersection de la ligne perpendiculaire à C_n -L_n passant par MR_n).

Le système décrit jusqu'à présent avec référence aux Fig. 1 à 9 est avantageusement complété par un système complémentaire qui va être décrit avec référence aux Fig. 11 à 16 pour constituer un système global 22 illustré sur la Fig. 10, sur laquelle on a représenté, d'une part, l'ensemble 11a de la Fig. 2 avec indication des signaux V, DI, VL, C, SR et composite F (HP, SL, SC) envoyés par l'ensemble 11 (qui reçoit le signal vidéo numérique d'entrée S) à l'ensemble 22a (qui débite une sortie composite ZH).

La ligne d'assemblage $Z - Z_1$ entre les ensembles 11a et 22a est illustrée sur les Fig. 2, 10 et 11, les sorties de l'ensemble 11a étant connectées aux entrées de l'ensemble 22a suivant $Z - Z_1$ pour transmettre les signaux précités.

La sortie de l'unité 22a et donc du dispositif global 22 est constituée par un signal composite ZH fournissant l'information désirée sur la zone en déplacement relatif de la scène 13a surveillée par la caméra vidéo 13.

L'unité complémentaire 22a, connectée aux sorties de l'unité 11a, est illustrée sur la Fig. 11 sous la forme de blocs fonctionnels, en rappelant que cette figure se raccorde suivant la ligne $Z - Z_1$ (à sa partie supérieure) à la Fig. 2 suivant la ligne $Z - Z_1$ (à la partie inférieure de cette dernière figure).

L'unité de la Fig. 11 constitue essentiellement un dispositif de formation et d'utilisation d'histogrammes et elle comporte :

- une micro-ligne bus 23 qui véhicule un certain nombre de signaux numériques, détaillés ci-après ;
- six blocs ou unités de formation et de traitement d'histogrammes, référencés 24, 25, 26, 27, 28, 29, respectivement pour les histogrammes de signaux vidéo numériques retardés SR, de vitesses V, de directions orientées (en code de Freeman) D1, de constantes de temps CO, de premiers axes $x(m)$ et de seconds axes $y(m)$;
- six blocs ou unités de combinaison linéaire 30, 31, 32, 33, 34, 35 associant leurs entrées en provenance du bus 23 pour former chacune un signal de validation V_1, V_2, V_3, V_4, V_5, V_6 pour les six blocs 24, 25, 26, 27, 28, 29 respectivement, 30 étant associée à 24, 31 étant associé à 25, 32 étant associé à 26, 33 étant associé à 27, 34 étant associé à 28 et 35 étant associé à 29 ;
- un bloc ou unité de zone en mouvement 36 coordonnant les sorties des blocs 28 et 29 pour les axes $x(m)$ et $y(m)$; et
- un bloc ou unité de changement de repère 37 recevant des signaux $x(m)_0$ et $y(m)_0$ d'orientation des axes $x(m)$ et $y(m)$, ainsi que des signaux d'horloge de pixel HP, de séquence ligne SL et de séquence colonne SC (ces trois signaux étant groupés dans le faisceau F des Fig. 2, 4, 10 et 11) en provenance de l'unité 19 des Fig. 2 et 4, et formant les signaux $x(m)_1$ et $y(m)_1$ envoyés respectivement aux unités 28 et 29.

Les unités 24, 25, 26, et 27 débitent chacune sa sortie SSR, SV, SDI, SDO sur le bus 23, alors que les unités 28 et 29 débitent chacune un signal $x(m)_2$, $y(m)_2$ sur une des deux entrées de l'unité de zone en mouvement 36, qui combine ces deux signaux en provenance des unités 28 et 29 et débite en sortie un signal composite $xy(m)$ sur le bus 23.

Le fonctionnement de chacune des unités 24 à 29 de formation et de traitement d'histogramme, qui reçoit en entrée, d'une part, un signal de validation V_1, V_2, V_3, V_4, V_5 ou V_6 de l'unité de combinaison linéaire 30 à 35 associée et, d'autre part, un signal SR, V, DI, CO, $x(m)_1$ ou $y(m)_1$ à traiter, est le même et c'est pourquoi, on exposera le fonctionnement d'une seule de ces unités, à savoir l'unité 25 concernant la formation et le traitement des histogrammes de vitesse V, prise à titre d'exemple ; seule la variable traitée est différente pour les autres unités analogues 24, 26, 27, 28 et 29, en notant que les différants signaux d'entrée pour les six unités 24 à 29 sont tous des signaux numériques, ce qui permet une analogie de structure et de fonctionnement des ces six unités.

Sur la Fig. 12, on a représenté schématiquement, chacun par son enveloppe, les histogrammes 38 et 39, respectivement en x et en y (axes de coordonnées cartésiennes de la matrice 21 à 17 x 17 éléments de la Fig. 4), des vitesses V de la zone en déplacement (sur la Fig. 14, on a indiqué les éléments tels que C_1, C_2 de l'histogramme à enveloppe 38) ; x_M et y_M

représentent les coordonnées en x et en y des maxima des deux enveloppes 38 et 39, respectivement, tandis que l_a et l_b, pour l'axe des x et l_c et l_d, pour l'axe des y représentent les limites du domaine des vitesses significatives ou intéressantes, l_a et l_c étant les limites inférieures et l_b et l_d les limites supérieures des portions significatives des histogrammes.

Les droites verticales L_a et L_b, d'abscisses l_a et l_b, et les droites horizontales L_c et L_d, d'ordonnées l_c et l_d, forment un rectangle qui encadre la zone hachurée 40 de vitesses significatives (pour l'ensemble des directions x et y), quelques microzones 41, de vitesses plus faibles et dont on ne tiendra pas compte, existant au voisinage de la zone principale 40.

Il suffit donc de repérer les coordonnées des quatre limites l_a, l_b, l_c, l_d et des deux maximas x_M et y_M pour caractériser la zone de plus grande variation du paramètre mis en histogramme, la vitesse V dans le cas particulier considéré. C'est cette information concernant V que le bloc 25 débite en sortie en permanence sur le micro-bus 23.

De même les blocs similaires 24, 26 et 27 débitent en sortie en permanence sur ce micro-bus 23 les informations concernant la zone de valeur maximale pour SR, DI et CO, respectivement.

Enfin les blocs similaires 28 et 29 débitent en sortie en permanence à l'unité 36 les informations concernant la zone de valeur maximale de $x(m)_1$ et $y(m)_1$ respectivement, l'unité 36 combinant ces informations d'abscisses et d'ordonnées, référencés $x(m)_2$ et $y(m)_2$ respectivement, en un signal composite $xy(m)$ qui est envoyé par la sortie de l'unité 36 sur le bus 23.

En définitive le bus 23 véhicule les informations concernant la zone de valeurs maximales de SR, V, DI, CO et $xy(m)$, c'est-à-dire $x(m)_1$ et $x(m)_2$, ce qui permet de déterminer l'existence d'une zone en déplacement dans la scène observée par la caméra vidéo 13, de localiser cette zone et d'en déterminer la vitesse et la direction (orientée) de déplacement.

Sur la Fig. 11 on a référencé ZH le signal composite de sortie disponible sur le bus 23. Les composantes précitées de ce signal ZH peuvent, notamment en ce qui concerne V et DI, c'est-à-dire la vitesse et la direction orientée de la zone en déplacement, être affichées sous forme numérique ou analogique, déclencher un signal lumineux et/ou sonore, notamment en cas de dépassement d'un seuil par la vitesse V, ou bien être transmis par câble, fibre optique ou voie hertzienne aux fins d'utilisation à distance, dans une unité de contrôle, telle que l'unité 10a de la Fig. 1, situé au voisinage ou à distance du dispositif 11 selon l'invention.

La portion des unités de la Fig. 12 au-dessus du bus 23 sert, à la suite du traitement point par point de toute une trame et l'association des points d'une trame pour en déduire une valeur globale externe, à déterminer l'existence et la localisation d'une zone de la scène observée en mouvement relatif et, en outre, s'il y a effectivement déplacement de cette zone, la vitesse et la

direction orientée de ce déplacement. Le repérage dans le plan d'observation de cette zone en mouvement relatif suivant deux directions x, y, qui peuvent ne pas être orthogonales (cas de l'application des Fig. 15 et 16 par exemple) est réalisé par la portion des unités de la Fig. 12 au-dessous du bus 23.

On va maintenant exposer plus en détail, avec référence aux Fig. 12, 13 et 14, la structure et le fonctionnement d'un bloc de formation et traitement d'histogramme, tel que 25, et de son bloc de combinaison linéaire associé, tel que 31.

Le bloc 25 (Fig. 13) comporte une portion 25a formant l'histogramme et une portion 25b constituant un classifieur pour l'histogramme, ces deux portions fonctionnent sous le contrôle d'un logiciel, matérialisé en une portion de circuit intégré 25c, qui réalise l'extraction des limites l_a, l_b, l_c, l_d de l'histogramme (Fig. 11).

Le classifieur 25b alimente, ainsi que les classifieurs des autres blocs de formation et traitement d'histogramme 24, 26, 27, 28, 29 (pour ces deux derniers à travers l'unité 36 de combinaison en $x(m)$ et $y(m)$), le bus 23 et, par l'intermédiaire de celui-ci, l'unité de combinaison linéaire 31 qui reçoit donc en parallèle des informations de tous les classifieurs des unités 24, 25, 26, 27, 28, 29 et qui, en fonction de ces informations, débite ou non le signal de validation V_2 à l'unité 25.

Grâce au logiciel de 25c, le classifieur 25b détermine les différentes classes (chacune comportant le même nombre de valeurs de la vitesse dans le cas envisagé) qui définiront l'enveloppe telle que 38 ou 39 (Fig. 12).

Sur la Fig. 14, on a illustré, suivant l'axe des x, les classes successives C_1, C_2 ... C_{n-1}, C_n et leur enveloppe 38 pour la vitesse V qui sont déterminées dans le classifieur 25b.

Les Fig. 15 et 16 illustrent le rôle des histogrammes pour $x(m)$ et $y(m)$ formés par les unités 28 et 29 et combinés dans l'unité 36 afin d'obtenir une pente.

A titre d'exemple on a considéré le cas de l'observation d'une route au moyen d'une caméra vidéo à sortie numérique embarquée dans un véhicule avec son dispositif selon l'invention associé.

Sur la Fig. 15 a représenté les deux bords de gauche B_g et de droite B_d d'une route R, ainsi que les pentes numérotées du 0 à 7 par exemple (suivant une convention autre que le code de Freeman) de la projection P_x en $x(m)$ réalisée par l'unité 28 et de la projection P_y en $y(m)$ réalisée par l'unité 29.

Pour assurer la précision la meilleure en ce qui concerne le bord de droite B_d, c'est-à-dire une sensibilité maximale des indications, par exemple de vitesse, concernant ce bord, il faut que la projection P_x ait la pente la plus voisine de la pente optimale P_0 qui est perpendiculaire à B_d, en

particulier la pente S dans la représentation de la Fig. 15. La valeur maximale de l'histogramme de vitesses sera donc obtenue pour la pente S déterminée par l'unité 28 (Fig. 11).

Le même raisonnement s'applique au bord gauche B_g en ce qui concerne la pente de la projection P_y et donc du rôle de l'unité 29.

L'unité 36 de combinaison des deux pentes optimales fournit l'information d'optimalisation pour l'ensemble des deux bords B_d et B_g.

La Fig. 16 illustre l'application de la détermination de la pente optimale P_o de la projection P_x, prise à titre d'exemple, pour assurer une conduite correcte d'un véhicule Vh en Europe continentale à conduite à droite de la route (cas a), au Royaume-Uni à conduite à gauche de la route (cas b) et enfin d'un avion Va pour un atterrissage correct au milieu de la piste d'un aéroport (cas c).

Par conséquent, pour aider la conduite d'un véhicule terrestre (auto, camion) sur une route ou d'un véhicule aérien (avion, navette spatiale) au voisinage d'une piste d'aéroport, le dispositif selon l'invention comporte en outre des moyens de représentation des bords droit Bd et gauche Bg de la route, respectivement de la piste, et des moyens pour orienter au moins un des axes, à pente variable, de coordonnées pour qu'il se maintienne sensiblement orthogonal au bord correspondant (position P_o).

Jusqu'à présent on a surtout décrit, les moyens pour déterminer une zone en déplacement effectif dans un environnement sensiblement immobile, grâce au repérage d'une région dans laquelle $DP \approx 1$. Pour déterminer une zone en repos dans un environnement dans l'ensemble en mouvement (cas de l'arrêt d'un véhicule en panne ou d'une collision sur une autoroute par exemple), c'est au contraire les zones dans lesquelles $DP \approx 0$ qui doivent être localisées par rapport à l'environnement pour lequel $DP \approx 1$. Bien entendu dans ce cas, les vitesses seront nulles dans une zone et la notion de direction n'a pas de signification. Les calculs dans 17a seront donc différents.

Le dispositif, selon l'invention, s'il ne doit servir qu'à déterminer une zone immobile peut être simplifié avec suppression des unités ou blocs traitant les vitesses et les directions, notamment les blocs 25, 26, 31, 32, et réduction du nombre des sorties de l'unité 11a et des entrées de l'unité 17.

Dans l'unité d'exploitation 10a du dispositif selon l'invention, on peut prévoir des moyens d'affichage, sur écran de moniteur, des histogrammes et/ou des valeurs des signaux DP ou CO.

Quant au signal SR, c'est-à-dire le signal vidéo numérique retardé, il est en général appliqué à un téléviseur ou moniteur 10 afin de visualiser, localement ou à distance sur l'écran de celui-ci, ce signal à l'instant où un mouvement relatif est signalé afin de vérifier la nature de ce

mouvement relatif. Il suffira donc d'observer l'écran du téléviseur ou moniteur 10 uniquement lorsqu'une zone en déplacement relatif a été signalée, par une alarme visuelle et/ou sonore par exemple.

On peut avantageusement faciliter le repérage d'une zone en mouvement relatif sur l'écran en imposant des couleurs arbitraires au signal vidéo numérique retardé SR, chaque couleur ou nuance de couleur représentant une vitesse et/ou une direction du mouvement.

Les différentes unités qui viennent d'être décrites avec référence aux Fig. 2, 3, 4, 11 et 13 sont réalisables chacune par des circuits électroniques de type connu, notamment des microprocesseurs effectuant des calculs et/ou des comparaisons ou utilisant : des signaux de balayage ; des mémoires ; des unités à retards ; des registres à décalage ; des unités formant des histogrammes linéaires et associant dans le plan de tels histogrammes ; des microbus.

La combinaison de ces circuits électroniques en unités 11a et 22a séparées ou en un ensemble 22 constitué par 11a et 22a peut être réalisée en deux circuits intégrés ou en un seul circuit intégré de très petite dimension, de l'ordre de 10 mm x 10 mm par exemple en technologie 0,7 µm, l'ensemble des deux unités intégrées interconnectées ou le circuit intégré unique étant connecté par son entrée à la sortie vidéo numérique d'une caméra vidéo ou autre dispositif d'observation et par leurs sorties à un ou plusieurs dispositifs d'utilisation sur place ou à distance. Dans une variante si on met en oeuvre seulement le dispositif simplifié de l'unité 11a, celui-ci, de préférence sous la forme d'un circuit intégré unique, est disposé entre ladite sortie numérique et un ou plusieurs dispositifs d'utilisation sur place ou à distance.

On va donner maintenant à titre d'exemples, nullement limitatifs, des applications supplémentaires d'un dispositif selon l'invention, notamment du dispositif selon les Fig. 1 et 10, c'est-à-dire selon l'ensemble des Fig. 2 et 11 raccordées suivant la ligne Z - Z'.

Une première application supplémentaire, illustrée sur la Fig. 17, est constituée par le cadrage automatique d'une personne se déplaçant dans une pièce, par exemple dans le cadre d'une vidéoconférence. Le cadrage automatique permet d'éliminer les mouvements de la personne se déplaçant, ce qui accroît la définition de l'image de cette personne observée par une caméra vidéo à sortie numérique et aussi, en cas de compression du signal vidéo numérique, simplifie cette compression.

Sur la Fig. 17, on retrouve la caméra vidéo 13 qui observe la personne P pouvant se déplacer. Le signal vidéo numérique S de la caméra vidéo est non seulement transmis par câble, fibre optique ou voie hertzienne à un écran de téléviseur ou moniteur 10b, mais également reçu par le dispositif 11 selon l'invention, dont la sortie ZH agit sur une unité 42 qui, en réponse aux signaux reçus du dispositif 11 concernant la localisation et le déplacement de la personne P,

commande les moteurs 43 de la caméra 13 pour diriger l'axe optique de la caméra vers la personne, notamment son visage F, en fonction de la localisation de son mouvement et de la vitesse et de la direction de celui-ci et éventuellement pour agir sur le réglage du zoom, de la distance focale et/ou de la mise au point de la caméra en cas d'avance ou du recul de la personne P.

On peut aussi commander le suivi par au moins un spot lumineux d'un personnage (acteur, chanteur) sur une scène, le dispositif selon l'invention centrant le personnage au centre de l'image par déplacement de l'orientation de la caméra vidéo et commandant la direction du spot, par exemple par un ou deux miroirs orientables, pour chaque spot.

Une autre application, donnée à titre d'exemple, du dispositif selon invention est illustrée sur la Fig. 18 sur laquelle on retrouve la caméra 13 ou autre dispositif d'observation qui débite un signal vidéo numérique S dans un dispositif 11 selon invention. Dans cette application, la caméra 13 surveille une portion d'autoroute afin de détecter l'arrêt inopiné d'un véhicule, en particulier sur la bande d'arrêt d'urgence, ou un arrêt de voiture à la suite d'une collision.

Il s'agit donc dans ce cas de déterminer l'immobilité d'un objet (le véhicule) dans un environnement en déplacement (les autres véhicules), c'est-à-dire de localiser dans la matrice 21 à 17 x 17 éléments la zone dans laquelle $DP \approx 0$. Normalement la caméra 13 observe un flot des véhicules qui produisent des réponses $DP \approx 1$, avec des valeurs de vitesse et de direction de déplacement. Par contre, si un véhicule s'arrête, il est signalé par une réponse $DP \approx 0$ dans sa zone d'observation.

Une unité 44 recevant ZH, ainsi que SR, détecte l'apparition d'une zone dans laquelle $DP \approx 0$ dans ZH et débite en sortie un signal d'anomalie NL qui, d'une part, déclenche une alarme sonore et/ou lumineuse dans un dispositif 45 et, d'autre part, commande un commutateur 46 qui applique le signal vidéo S (ou plutôt le signal vidéo retardé SR) sur l'écran d'un téléviseur ou moniteur 10 qui permet au surveillant, alerté par l'alarme sonore et/ou visuelle, d'observer l'autoroute à l'instant où un véhicule s'arrête ou de collision afin de pouvoir prendre les mesures nécessaires, par exemple suivant la réaction de conducteur du véhicule qui s'est arrêté.

Grâce à l'invention, dans un poste de surveillance, un surveillant peut facilement contrôler un grand nombre de tronçons sur une autoroute (ou une route) dans chacun desquels est disposé une caméra 13 et un dispositif 11 selon l'invention, les sorties ZH et SR de chaque dispositif 11 étant transmises par câble, fibre optique ou voie hertzienne à un seul poste de surveillance où se trouve une unité commune 44 ; en effet l'observateur n'a besoin de regarder l'écran de 10 qu'en cas d'incident ou accident, signalé par l'alarme, et il est rare que plusieurs incidents et/ou accidents en des emplacements différents se produisent simultanément.

En particulier l'unité 44 peut comporter, en entrée, un commutateur de type rotatif (non représenté) envoyant successivement et cycliquement les signaux ZH (et SR) des différents ensembles 13 - 11 disposés le long de l'autoroute sur la portion de cette unité 44 produisant le signal NL.

Le même système permet de déceler non seulement, comme indiqué, l'arrêt, ou la collision entraînant un arrêt de véhicules, mais également un ralentissement du flot des véhicules (en cas de circulation très ralentie) par diminution de la vitesse dans les différentes zones dans lesquelles D = 1 et inversement un véhicule à vitesse excessive, la vitesse dans une zone observée dépassant alors la vitesse limitée autorisée.

Enfin sur la Fig. 19 on a illustré une autre application de l'invention, à savoir à l'interaction homme - machine, le déplacement de la main M, ou plus particulièrement des doigts DG, à l'intérieur d'une surface SF décomposée en rectangles par un système de coordonnées C_x et C_y.

Une caméra vidéo 13 à sortie numérique associée à un dispositif 11 selon l'invention, tel qu'illustré sur la Fig. 1, permet de reconnaître le mouvement d'une main M et des doigts DG et de l'utiliser pour contrôler un ordinateur (à la manière d'une «souris») ou pour commander certaines fonctions dans une machine. Par exemple l'ensemble 13 - 11 pourrait servir à un sourd-muet, utilisant le code standard du langage des sourds-muets à base de mouvements des mains, pour entrer en ordinateur des données alphanumériques, donc un texte, sans avoir à utiliser le clavier habituel ; cette opération peut, bien entendu être également effectuée par une personne parlante, ayant appris le langage des sourds-muets, afin d'entrer un texte en ordinateur sans avoir à utiliser un clavier. Un tel système n'est pas sensible aux écarts temporels et n'exige pas une signalisation précise du début et de la fin du geste.

Les Fig. 20 et 21 illustrent schématiquement l'application de l'invention à la surveillance d'un conducteur automobile afin de signaler l'endormissement de celui-ci.

La caméra vidéo 13 est dans ce cas placée contre la portion de carrosserie à l'intérieur du véhicule, au-dessus du rétroviseur par exemple, et observe le conducteur.

L'opération préliminaire consiste à cadrer celui-ci comme dans le cas de l'application de la Fig. 17. Sur la Fig. 20, on a représenté symboliquement l'image IC du conducteur sur l'écran vidéo. On supprime d'abord les portions inutiles de droite et de gauche (hachures horizontales de l'image et on se limite ainsi à la portion centrale de l'image entre ces deux portions.

Ensuite dans cette portion centrale il suffit de surveiller la zone non hachurée AA de la Fig. 21 dans laquelle est cadrée la tête.

Les déplacements intéressants qui sont détectés par le dispositif selon l'invention sont constituées par les clignements des paupières du conducteur (signalés par des mouvements verticaux dans la zone AA) dont la cadence se modifie en préliminaire à l'endormissement. Si la cadence ou et la vitesse de ces clignements devient inférieure à un certain seuil, une alarme sonore est déclenchée et le conducteur est réveillé.

Sur la Fig. 22, on a indiqué un moyen pour pallier, lorsque cela est nécessaire, au nombre limité de bits p représentatifs de la constante de temps CO pour permettre de prendre en compte une plus large gamme de vitesses de déplacement.

A cet effet on prévoit l'utilisation du diagramme de Mallat (voir article de S. Mallat «A Theory for multiresolution signal decomposition» in IEEE Transactions on Pattern Analysis and Machine Intelligence, Juillet 1989 p. 674-693). qui consiste à décomposer successivement la totalité de l'image vidéo en moitié successives, repérées 1, 2, 3, 4, 5, 6, 7. On réalise ainsi une compression en ne traitant que des portions d'images. On peut ainsi avec $p = 4$, c'est-à-dire $2^p = 16$ déterminer une vitesse dans une gamme plus étendue.

Si au début, dans le cadre de l'image totale, le dispositif selon invention indique que la vitesse de l'objet (au sens large) mobile dépasse la vitesse maximale déterminable avec $2^p = 16$ pour la constante de temps, il suffit de passer successivement par les images partielles observées 1, 2, 3, 4,... jusqu'à ce que la vitesse de l'objet mobile ne dépasse pas ladite vitesse maximale dans le cadre de l'image partielle après compression.

Pour mettre en oeuvre la composition de Mallat par ondelettes, il suffit d'intercaler sur le schéma de la Fig. 1 une unité 13A (illustrée sur la Fig. 22) qui réalise cette compression du signal vidéo. Cette unité peut par exemple être constituée par le composant «ADV 601 Low Cost Multiformat Vidéo Codec» de la société américaine ANALOG DEVICES faisant l'objet de la notice «ADV 601 Preliminary Data Sheet» de Janvier 1996. Sur la Fig. 2, une telle unité de compression optionnelle 13a.

Enfin sur les Fig. 2 et 3, on a indiqué sur certains microbus le nombre de bits (1, 3 (dans le cas où $p = 3$), 8, 15) véhiculés, ce qui montre l'économie de taille possible pour les différents blocs fonctionnels ayant à traiter un nombre réduit de bits.

On voit en définitive que l'invention permet de détecter un mouvement relatif dans une scène observée par un dispositif optoélectronique, tel qu'une caméra vidéo, qui transmet la scène observée sous la forme d'un signal vidéo numérique constitué par une succession de trames, elles-mêmes constituées par une succession de lignes composées d'une succession de pixels, ce signal numérique étant analysé afin de repérer une zone en déplacement relatif, avec indication de la

vitesse et de la direction (orientée) de ce déplacement si la zone est en mouvement effectif par rapport à un environnement sensiblement immobile.

Etant donné que le dispositif selon l'invention détermine la direction orientée et la vitesse de déplacement d'un objet (au sens le plus large), on peut lui adjoindre des moyens pour déduire de ces deux parametres une position future de l'objet à un instant déterminé et des moyens pour orienter d'avance la caméra vidéo d'entrée 13 sur cette position future.

Il y a lieu de noter que les résultats obtenus avec le dispositif selon invention n'exigent nullement que la caméra soit immobile, ce qui permet d'embarquer la caméra et le dispositif associé sur un véhicule terrestre, aérien ou nautique (pour la mise en oeuvre du procédé illustré sur la Fig. 16 par exemple).

Après une très courte période d'initialisation égale à N (de l'ordre de moins de dix trames successives correspondantes), le dispositif selon l'invention détermine les paramètres du déplacement relatif instantanément après la fin de chaque trame ayant subi les traitements temporel et spatial, du fait de la récursivité des calculs selon l'invention.

On a décrit un mode de réalisation préféré du dispositif selon l'invention et quelques applications de celui-ci. Bien entendu ce mode de réalisation et ces applications ont été donnés à titre d'exemples non limitatifs et de nombreuses variantes et adaptations, qui apparaîtront facilement à l'homme de l'art, peuvent être envisagées sans sortir du cadre de l'invention tel que défini dans les revendications ci-après.

On pourrait par exemple mettre en oeuvre des moyens d'utilisation des signaux débités par l'unité 11 de la Fig. 2 autres que ceux illustrés sur la Fig. 11 sans s'écarter de l'invention.

Quant aux applications d'un dispositif selon l'invention, elles ne sont nullement limitées à celles données à titre d'exemple dans la description précitée. Ainsi un montage analogue à celui de la Fig. 17 peut être associé directement à un caméscope pour le stabiliser relativement aux déplacements causés par les mouvements involontaires de l'utilisateur.

On peut également utiliser un ou, de préférence, plusieurs dispositifs, selon l'invention, associés respectivement à un ou, de préférence, plusieurs caméscopes à sortie vidéo numérique placés dans une pièce d'immeuble pour constituer une «pièce intelligente» grâce à un tel système permettant de détecter et localiser la présence et le mouvement d'une ou plusieurs personnes dans la pièce, d'analyser ce mouvement, aux fins de sécurité, d'identification et/ou pour aider dans les tâches à accomplir, par exemple, ou à surveiller des enfants dans une autre pièce ou les clients d'un supermarché.

On a supposé essentiellement l'utilisation d'un signal vidéo à paires de trames entrelacées successives, notamment lors de la discussion de la capacité de la mémoire 16, avec traitement des

deux trames d'une paire dans le dispositif selon l'invention. Il est toutefois possible de n'utiliser qu'une trame sur deux (l'impaire par exemple) en réduisant la capacité de la mémoire, mais avec réduction de moitié environ de la vitesse d'obtention des informations désirées. Il est également possible de mettre en oeuvre une camera vidéo ou autre dispositif d'observation dont la sortie numérique ne comporte qu'une seule trame par image.

Il est possible, dans certaines applications d'associer, au dispositif selon l'invention, des capteurs spécialisés, par exemple un ou plusieurs capteurs d'accélération, pour pouvoir traiter des parametres supplémentaires du déplacement.

Bien entendu l'invention n'est pas limitée aux modes de réalisation particulières ni aux applications décrits, mais elle en embrasse toutes les variantes et modifications entrant dans la définition générale de l'invention.

REVENDICATIONS

1. Procédé, en temps réel, pour le repérage et la localisation d'une zone en mouvement relatif dans une scène observée par un système d'observation à sortie constituée par un signal vidéo numérique du type comportant une succession de trames correspondantes, chacune composée d'une succession de lignes, composées chacune d'une succession de pixels, ainsi que pour la détermination de la vitesse et de la direction orientée du déplacement, ledit procédé étant caractérisé en ce qu'il consiste à effectuer sur le signal vidéo numérique de sortie successivement :

- un traitement de lissage dudit signal vidéo numérique de sortie mettant en oeuvre une constante de temps numérique dont la valeur numérique peut être modifiée pour chacun des pixels dudit signal de sortie, indépendamment pour chacun d'entre eux ;
- une mise en mémoire d'une trame dudit signal de sortie après le lissage, d'une part, et de la constante de temps de lissage associée à la dite trame, d'autre part ;
- un traitement temporel consistant, pour chaque position de pixel, à déterminer l'existence, d'une part, et l'amplitude, d'autre part, d'une variation significative de l'amplitude du signal de pixel entre la trame actuelle et la trame juste antérieure lissée, mise en mémoire, et à générer deux signaux numériques, le premier signal étant un signal binaire ou monobit à deux valeurs possibles dont l'une représente l'existence d'une telle variation significative et l'autre l'absence d'une telle variation entre deux trames successives, la valeur dudit signal binaire modifiant la valeur mémorisée de ladite constante de temps afin de la diminuer si ledit signal représente une variation significative et afin de l'augmenter si ce signal respectivement ne représente pas une telle variation , la diminution ou l'augmentation étant réalisée d'une manière quantifiée, tandis que le second signal numérique, dit d'amplitude, est un signal multibits à nombre limité de bits, quantifiant l'amplitude de cette variation ; et
- un traitement spatial consistant, pour chaque trame de signal vidéo numérique d'entrée,
  - à répartir, afin de caractériser les valeurs des pixels, en une matrice à nombre de lignes et nombre de colonnes réduits, par rapport respectivement au nombre de lignes et au nombre de pixels par ligne dans le signal vidéo, uniquement les valeurs, au même instant d'observation, pour une fraction des pixels d'une trame - fraction qui défile par balayage à travers ladite matrice pendant la durée d'une trame - dudit signal binaire, d'une part, et dudit signal numérique d'amplitude, d'autre part,
  - à déterminer, dans cette double représentation matricielle instantanée, une zone particularisée, dans laquelle à la fois ledit signal binaire a la valeur recherchée

représentant la présence, ou respectivement l'absence, de variation significative et ledit signal numérique d'amplitude varie, ou respectivement ne varie pas, d'une valeur significative entre pixels voisins dans la matrice suivant une direction orientée à partir d'un pixel d'origine, et ceci pour une même portion de trame, donc à un même instant d'observation, et

- à engendrer des signaux représentatifs de l'existence et de la localisation de la zone en déplacement relatif, ainsi que la vitesse relative intertrames et la direction orientée de ce déplacement, s'il existe, relativement à son environnement, à partir de la répartition matricielle instantanée de ces deux signaux numériques, binaire et d'amplitude

2. Procédé selon la revendication 1, caractérisé en ce qu'il consiste, en outre,

- à former les histogrammes des valeurs des signaux répartis matriciellement d'une part, et les histogrammes des valeurs des inclinaisons de deux axes, à pente variable, de coordonnées dans un plan, d'autre part,
- à repérer, dans chaque histogramme formé, un domaine de variation significative de la valeur traitée et
- à déduire, de chaque domaine repéré, l'existence et la localisation, ainsi que la vitesse et la direction orientée, d'une zone en mouvement relatif.

3. Procédé selon la revendication 1 ou 2, caractérisé en ce que ladite matrice est une matrice carrée à même nombre impair $(2I + 1)$ de lignes et de colonnes, et que l'on considère les matrices emboîtées de 3×3, 5×5, 7×7, ... $(2I + 1) \times (2I + 1)$ éléments centrées sur le centre de cette matrice carrée afin de déterminer la matrice emboîtée de plus petite taille dans laquelle ledit signal numérique varie dans une direction orientée à partir dudit centre, la valeur dudit signal binaire représentant un dépassement du seuil selon cette direction.

4. Procédé selon la revendication 1 ou 2, caractérisé en ce que ladite matrice est une matrice hexagonale et que l'on considère les matrices hexagonales emboîtées de taille croissante centrées sur le centre de cette matrice hexagonale afin de déterminer la matrice emboîtée de plus petite taille dans laquelle ledit signal numérique varie dans une direction orientée à partir dudit sommet, la valeur dudit signal binaire représentant un dépassement du seuil selon cette direction.

5. Procédé selon les revendications 1 ou 2, caractérisé en ce qui ladite matrice est une matrice en L renversé à une seule ligne et une seule colonne et en ce que l'on considère les matrices emboîtées de 3×3 pixels, 5×5 pixels, 7×7 pixels... $(2I + 1) \times (2I + 1)$ pixels, pour la ligne et la colonne uniques, afin de déterminer la matrice de plus petite taille dans laquelle le

signal varie dans une direction orientée, à savoir la ligne de plus grande pente à quantification constante.

6. Procédé l'une quelconque des revendications précédentes, caractérisé en ce que ladite constante de temps est de la forme 2^p, p étant un nombre inférieur à 16, pouvant donc être exprimé par pas plus de 4 bits, la diminution ou l'augmentation de la constante de temps étant réalisée par la soustraction ou l'addition d'une unité à p.

7. Procédé selon la revendication 6, caractérisé en ce que l'on considère des portions successives décroissantes de trames complètes suivant l'algorithme temps - échelle de Mallat et sélectionne la plus grande de ces portions, qui donne des indications de déplacement, vitesse et orientation, compatible avec la valeur de p.

8. Dispositif fonctionnant en temps réel, pour le repérage et la localisation d'une zone en mouvement relatif dans une scène observée par un système d'observation à sortie constituée par un signal vidéo numérique, du type comportant une succession de trames correspondantes, des lignes successives dans chaque trame correspondante et des pixels successifs dans chaque ligne, ainsi que pour la détermination de la vitesse et de la direction orienté de ce déplacement, par mise en oeuvre du procédé selon la revendication 1, le dispositif recevant, en entrée, ledit signal vidéo numérique de sortie, caractérisé en ce qu'il comporte en combinaison :

- des moyens de lissage (15) dudit signal vidéo numérique de sortie mettant en oeuvre une constante de temps numérique (CO) dont la valeur numérique peut être modifiée pour chacun des pixels dudit signal de sortie, indépendamment pour chacun d'entre eux ;
- des moyens de mémorisation (16) d'une trame dudit signal de sortie après lissage (LI), d'une part, et de la constante de temps de lissage (CI) associée à ladite trame, d'autre part ;
- une unité de traitement temporel (15) pour analyser les variations temporelles de l'amplitude du signal de pixel, pour une même position de pixel, entre la trame actuelle et la trame juste antérieure lissée, mise en mémoire, dudit signal vidéo numérique, ladite unité comportant, en association avec une mémoire apte à recevoir, stocker et restituer des informations relatives à la trame correspondante précédente lissée, des moyens de comparaison (15a) pour déterminer, pour chaque position de pixel dans la trame du signal vidéo entrant, si la valeur absolue de la différence entre le signal de pixel actuel et une valeur représentative du signal de pixel, pour la même position de pixel, dans la trame antérieure, valeur représentative stockée dans ladite mémoire, dépasse ou non un seuil, en générant un signal binaire ou monobit (DP) à deux valeurs, dont l'une représente l'existence d'un dépassement et dont l'autre représente l'absence d'un dépassement, et des moyens de calcul (15c) aptes à déterminer un signal numérique d'amplitude multibits, à nombre réduit

de bits, dont la valeur est fonction de l'amplitude de la variation de la valeur du même pixel entre la trame actuelle et la trame juste antérieure lissée, mise en mémoire, du signal vidéo numérique ; et

- un ensemble de traitement spatial (11), dont les entrées reçoivent, de l'unité de traitement temporel, lesdits signaux binaire et numérique d'amplitude successifs pour les pixels d'une même trame, ladite unité comportant des moyens aptes à caractériser les valeurs d'amplitude des pixels, ces moyens répartissant, suivant une matrice (21) à nombre de lignes et nombre de colonnes réduits par rapport respectivement au nombre de lignes et au nombre de pixels par ligne dans une trame dudit signal vidéo numérique, afin de caractériser les valeurs des pixels uniquement lesdits signaux binaire et numérique d'amplitude relatif à un même instant, c'est-à-dire à une même trame, celle-ci défilant par balayage à travers ladite matrice pendant la durée d'une trame, des moyens de repérage pour déterminer, dans ladite matrice, une zone de pixels dans laquelle, à cet instant, le signal binaire a la valeur recherchée et des moyens pour déterminer, dans ladite matrice, une zone de pixels dans laquelle, à ce même instant, le signal numérique d'amplitude varie d'une quantité significative entre pixels voisins, et des moyens qui, en réponse aux indications des deux derniers moyens précédents, génèrent des signaux représentatifs de cette zone de pixels, donc de l'existence et de la localisation d'une zone en mouvement relatif dans la scène observée, ainsi que de la vitesse relative intertrames et de la direction orientée de cette zone lorsqu'elle se déplace effectivement relativement à son environnement.

9. Dispositif selon la revendication 7, caractérisé en ce que l'ensemble de traitement spatial (17, 18) comporte des premiers moyens de retard (*r*) en cascade dont chacun impose un retard égal à l'écart temporel entre deux lignes successives et des seconds moyens de retard (*d*) en cascade pour chaque ligne imposant chacun un retard égal à l'écart temporel entre deux pixels successifs d'une ligne, les sorties de chacun desdits seconds moyens de retard (*d*) et l'entrée de la cascade desdits seconds moyens de retard (*d*) de chaque ligne débitant à un instant donné les valeurs dudit signal binaire et dudit signal numérique d'amplitude, à un même instant, vers lesdits moyens de repérage (17a).

10. Dispositif selon la revendication 8, ou 9, caractérisé en ce qu'il comporte également des moyens (24 - 29) formant les histogrammes des valeurs de sortie de ladite unité de traitement spatial, ainsi que les histogrammes des inclinaisons de deux axes, à pente variable, de coordonnées dans un plan, des moyens pour repérer, dans chaque histogramme, un domaine de variation significative de la valeur traitée, afin de valider ce domaine en leur sortie et de déduire, pour l'ensemble des histogrammes, des signaux de sortie qui signalent et localisent

une zone de la scène observée en mouvement relatif, si elle existe, ainsi que la vitesse et la direction orientée de ce mouvement si ladite zone se déplace effectivement par rapport à son environnement.

11. Dispositif selon la revendication 8, 9, ou 10, caractérisé en ce lesdits moyens de lissage (15c, 15d) comportent une entrée qui reçoit ledit signal vidéo numérique S(PI) et calculent, pour les pixels successifs d'une trame de ce signal vidéo, un signal lissé (LO), dans lequel les variations temporelles du signal vidéo numérique d'entrée sont diminuées, par mise en oeuvre d'un signal de seuil (N) reçu sur une autre entrée et d'une constante de temps (CO) relative à chaque position de pixel d'une trame, dont la valeur est successivement modifiée afin que le lissage conserve, tout en la réduisant, la tendance de variation du signal vidéo numérique entrant, ces moyens de lissage coopérant avec l'unité de mémoire (16) qui reçoit, stocke et restitue les valeurs actualisées, pour chaque portion de pixel d'une trame, du signal lissé et de ladite constante de temps et débitent, sur leurs sorties, au moins la succession, pour chaque position de pixel, des valeurs de la constante de temps actualisée et des valeurs d'un signal binaire de dépassement ou non-dépassement dudit seuil par la valeur absolue de la différence entre la valeur du pixel et sa valeur lissée.

12. Dispositif selon l'une quelconque des revendications 8 à 11, caractérisé en ce que ladite unité de traitement spatial (17, 18) réalisant la répartition matricielle, par lignes et par colonnes en nombre réduit, des sorties desdits moyens de lissage, à savoir des valeurs successives de la constante de temps (CO) et dudit signal binaire, comporte des moyens de repérage (17a) pour repérer, dans ladite répartition matricielle, une zone de pixels dans laquelle, à la fois, soit la valeur dudit signal binaire correspond à un dépassement de seuil et ladite constante de temps varie entre pixels voisins d'une valeur significative dans une direction, et pour produire des signaux de sortie indiquant la localisation de ladite zone et la vitesse et la direction orientée du déplacement dans ladite zone, soit la valeur dudit signal binaire correspond à un non-dépassement du seuil et ladite constante de temps ne varie pas entre pixels voisins.

13. Application du dispositif selon la revendication 10, 11 ou 12, pour aider la conduite d'un véhicule terrestre sur une route, respectivement d'un véhicule aérien au voisinage d'une piste d'aéroport, caractérisé en ce que le dispositif comporte en outre des moyens de représentation des bords droit Bd et gauche Bg de la route, respectivement de la piste, et des moyens pour orienter au moins un des axes, à pente variable, de coordonnées pour qu'il se maintienne sensiblement orthogonal au bord correpondant (position P_0).

REVENDICATIONS MODIFIEES
[reçues par le Bureau International le 16 janvier 1998 (16.01.98);
revendications 1-13 remplacées par les revendications 1-14 modifiées (10 pages)]

1. Procédé, opérant en temps réel, pour le repérage et la localisation d'une zone en mouvement relatif dans une scène observée par un système d'observation à sortie constituée par un signal vidéo numérique du type comportant, d'une manière classique, une succession de trames correspondantes de même nature, chacune composée d'une succession de lignes, composées chacune d'une succession de pixels, ainsi que pour la détermination de la vitesse et de la direction orientée du mouvement de déplacement relatif, ledit procédé

consistant à effectuer sur ledit signal vidéo numérique successivement

- un traitement temporel qui consiste, pour chaque position de pixel, à déterminer la différence entre l'amplitude du signal de pixel dans la trame actuelle et l'amplitude d'un signal de pixel représentatif de ses valeurs dans les trames antérieures, et

- un traitement spatial, qui consiste, pour chaque trame dudit signal vidéo numérique, à répartir, en une matrice à nombre de lignes et nombre de colonnes réduits par rapport respectivement au nombre de lignes et au nombre de pixels par ligne dans ledit signal vidéo numérique, uniquement les valeurs, au même instant d'observation, pour une fraction des pixels d'une trame - fraction qui défile par balayage à travers ladite matrice pendant la durée d'une trame -, des signaux représentatifs des variations des valeurs de pixels et à déduire, à partir de l'ensemble des représentations matricielles relatives à un instant d'observation donné, c'est-à-dire à une trame donnée dudit signal vidéo ou numérique, la localisation d'une zone éventuelle en mouvement relatif et l'estimation de ce mouvement s'il existe,

et étant caractérisée en ce que

- il comporte un double traitement préliminaire consistant en

- un lissage adaptatif dudit signal vidéo numérique en mettant en oeuvre une constante de temps numérique dont la valeur numérique, pour chacun des pixels dudit signal vidéo numérique et indépendamment pour chacun d'eux, est modifiée en réponse à la variation ou non variation temporelle de la valeur du pixel, et
- à une mise en mémoire d'une trame entière dudit signal de sortie après lissage, d'une part, et de ladite constante de temps pour chaque position de pixel de ladite trame, d'autre part;

- ledit traitement temporel consiste, pour chaque position de pixel, à déterminer l'existence, d'une part, et l'amplitude, d'autre part, d'une variation significative éventuelle de l'amplitude du signal de pixel entre la trame actuelle et la trame juste antérieure lissée, mise en mémoire, et à générer deux signaux numériques , le premier étant un signal binaire, donc monobit, à deux valeurs possibles dont la première représente l'existence d'une telle variation significative et la seconde l'absence d'une telle variation significative entre les deux dites trames actuelle et juste antérieure lissée, la valeur dudit signal binaire modifiant la valeur mémorisée de ladite constante de temps pour le pixel en cause afin respectivement de la diminuer si ledit signal binaire a ladite première valeur et de l'augmenter si ledit signal binaire a ladite seconde valeur, la diminution ou l'augmentation étant réalisée d'une manière quantifiée, tandis que le second signal numérique, dit d'amplitude, est un signal multibits, à nombre limité de bits, quantifiant l'amplitude, de cette variation; et

- ledit traitement spatial consiste, pour chaque trame dudit signal vidéo de sortie,
  - à répartir, en ladite matrice et pendant la durée d'une trame, ledit signal binaire, d'une part, et ledit signal numérique d'amplitude, d'autre part,

à déterminer, dans cette double représentation matricielle instantanée dudit signal binaire et dudit signal numérique d'amplitude pour le même pixel, une zone particularisée, dans laquelle à la fois ledit signal binaire a la valeur recherchée représentant la présence, ou respectivement l'absence, de variation significative et ledit signal numérique d'amplitude varie, ou respectivement ne varie pas, d'une valeur significative entre pixels voisins dans ladite matrice suivant une direction orientée à partir d'un pixel d'origine, et ceci pour une même portion de trame, donc à un même instant d'observation, et

- à engendrer des signaux représentatifs de l'existence et de la localisation de la zone de déplacement relatif éventuel, de la valeur de la vitesse relative inter-trames et de la direction orientée parmi plusieurs directions orientées possibles dans le plan d'observation, de ce déplacement relatif éventuel, dans son environnement, à partir de ladite répartition matricielle instantanée de ces deux signaux, binaire et d'amplitude.

2. Procédé selon la revendication 1, caractérisé en ce qu'il consiste, en outre:

- à former les histogrammes des valeurs des signaux répartis matriciellement, d'une part, et les histogrammes de valeurs des inclinaisons de deux axes, à pente variable, de coordonnées dans un plan, d'autre part,
- à repérer, dans chaque histogramme formé, un domaine de variation significative de la valeur traitée et
- à déduire, de chaque domaine repéré, l'existence et la localisation, ainsi que la vitesse et la direction orientée, d'une zone en mouvement relatif.

3. Procédé selon la revendication 1 ou 2, caractérisé en ce que ladite matrice est une matrice carrée à même nombre impair (2/ + 1) de

lignes et de colonnes, et que l'on considère les matrices emboîtées de 3 x 3, 5 x 5, 7 x 7, ...(2*l* + 1) x (2*l* + 1) éléments centrées sur le centre de cette matrice carrée afin de déterminer la matrice emboîtée de plus petite taille dans laquelle ledit signal numérique varie dans une direction orientée à partir dudit centre, la valeur dudit signal binaire représentant un dépassement du seuil selon cette direction.

4. Procédé selon la revendication 1 ou 2, caractérisé en ce que ladite matrice est une matrice hexagonale et que l'on considère les matrices hexagonales emboîtées de taille croissante centrées sur le centre de cette matrice hexagonale afin de déterminer la matrice emboîtée de plus petite taille dans laquelle ledit signal numérique varie dans une direction orientée à partir dudit sommet, la valeur dudit signal binaire représentant un dépassement du seuil selon cette direction.

5. Procédé selon les revendications 1 ou 2, caractérisé en ce que ladite matrice est une matrice en L renversé à une seule ligne et une seule colonne et en ce que l'on considère les matrices emboîtées de 3 x 3 pixels, 5 x 5 pixels, 7 x 7 pixels....(2*l* + 1) x (2*l* + 1) pixels, pour la ligne et la colonne uniques, afin de déterminer la matrice de plus petite taille dans laquelle le signal varie dans une direction orientée, à savoir la ligne de plus grande pente à quantification constante.

6. Procédé selon l'une quelconque des revendications précédentes, caractérisé en ce que ladite constante de temps est de la forme 2^p, *p* étant un nombre inférieur à 16, pouvant donc être exprimé par pas plus de 4 bits, la diminution ou l'augmentation de la constante de temps étant réalisée par la soustraction ou l'addition d'une unité à *p*.

7. Procédé selon la revendication 6, caractérisé en ce que l'on considère des portions successives décroissantes de trames complètes suivant l'algorithme temps - échelle de Mallat et sélectionne la plus

grande de ces portions, qui donne des indications de déplacement, vitesse et orientation, compatible avec la valeur de *p*.

8. Dispositif, fonctionnant en temps réel, pour le repérage et la localisation d'une zone en mouvement relatif dans une scène observée, ainsi que pour la détermination de la vitesse et de la direction orientée du mouvement de déplacement relatif, par mise en oeuvre du procédé selon la revendication 1,

comportant:

- un système d'observation à sortie constituée par un signal vidéo numérique du type comportant, d'une manière classique, une succession de trames correspondantes de même nature, chacune composée d'une succession de lignes, composées chacune d'une succession de pixels,

- un ensemble de traitement temporel déterminant, pour chaque position de pixel, la différence entre l'amplitude du signal de pixel dans la trame actuelle et l'amplitude d'un signal de pixel représentatif de ses valeurs dans les trames antérieures,

- un ensemble de traitement spatial qui, pour chaque trame dudit signal vidéo numérique, répartit en une matrice à nombre de lignes et nombre de colonnes réduits par rapport respectivement au nombre de lignes et au nombre de pixels par ligne dans ledit signal vidéo numérique, uniquement les valeurs, au même instant d'observation, pour une fraction des pixels d'une trame - fraction qui défile par balayage à travers ladite matrice pendant la durée d'une trame -, des signaux représentatifs des variations des valeurs de pixels, et qui déduit, à partir de l'ensemble des représentations matricielles relatives à un instant d'observation donné, c'est-à-dire à une trame donnée dudit signal vidéo numérique, la localisation d'une zone éventuelle en mouvement relatif et l'estimation de ce mouvement s'il existe,

et étant caractérisée en ce qu'il comporte en combinaison:

- des moyens de lissage adaptatif (15) dudit signal vidéo numérique en mettant en oeuvre une constante de temps numérique (CO) dont la valeur numérique, pour chacun des pixels dudit signal vidéo numérique et indépendamment pour chacun d'eux, est modifiée en réponse à la variation ou non variation temporelle de la valeur du pixel;

- des moyens de mémorisation (16) pour une trame entière dudit signal de sortie après lissage (LI), d'une part, et de ladite constante de temps pour chaque position de pixel de ladite trame, d'autre part;

- dans l'ensemble de traitement temporel (15),

- o des moyens de comparaison (15) pour déterminer l'existence, d'une part, et l'amplitude, d'autre part, d'une variation significative éventuelle de l'amplitude du signal de pixel entre la trame actuelle et la trame juste antérieure lissée, mise en mémoire, et pour générer un premier signal numérique (DP) , qui est un signal binaire, donc monobit, à deux valeurs possibles dont la première représente l'existence d'une telle variation significative et la seconde l'absence d'une telle variation significative entre les deux dites trames actuelle et juste antérieure lissée, la valeur dudit signal binaire modifiant la valeur mémorisée de ladite constante de temps pour le pixel en cause afin respectivement de la diminuer si ledit signal binaire a ladite première valeur et de l'augmenter si ledit signal binaire a ladite seconde valeur, la diminution ou l'augmentation étant réalisée d'une manière quantifiée, et
- o des moyens de calcul (15c) aptes à générer un second signal numérique (CO), dit d'amplitude, qui est un signal multibits, à nombre limité de bits, quantifiant l'amplitude de cette variation; et

- dans l'ensemble de traitement spatial (11), dont les entrées reçoivent de ladite unité de traitement lesdits signaux numériques, binaire et d'amplitude, successifs pour les pixels d'une même trame:

- des moyens pour répartir, en ladite matrice (21) et pendant la durée d'une trame, ledit signal binaire, d'une part, et ledit signal numérique d'amplitude, d'autre part,
- des moyens pour déterminer, dans cette double représentation matricielle instantanée dudit signal binaire et dudit signal numérique d'amplitude pour le même pixel, une zone particularisée, dans laquelle à la fois ledit signal binaire a la valeur recherchée représentant la présence, ou respectivement l'absence, de variation significative et ledit signal numérique d'amplitude varie, ou respectivement ne varie pas, d'une valeur significative entre pixels voisins dans ladite matrice suivant une direction orientée à partir d'un pixel d'origine, et ceci pour une même portion de trame, donc à un même instant d'observation, et
- des moyens pour engendrer des signaux représentatifs de l'existence et de la localisation de la zone de déplacement relatif éventuel, de la valeur de la vitesse relative inter-trames et de la direction orientée parmi plusieurs directions orientées possibles dans le plan d'observation, de ce déplacement relatif éventuel, dans son environnement, à partir de ladite répartition matricielle instantanée de ces deux signaux, binaire et d'amplitude.

9. Dispositif selon la revendication 8, caractérisé en ce que l'ensemble de traitement spatial (17, 18) comporte des premiers moyens de retard (r) en cascade dont chacun impose un retard égal à l'écart temporel entre deux lignes successives et des seconds moyens de retard (d) en cascade pour chaque ligne imposant chacun un retard égal à l'écart temporel entre deux pixels successifs d'une ligne, les sorties de chacun desdits seconds moyens de retard (d) et l'entrée de la cascade desdits seconds moyens de retard (d) de chaque ligne débitant à un instant donné les valeurs dudit signal binaire et dudit signal numérique d'amplitude, à un même instant, vers lesdits moyens de repérage (17a).

10. Dispositif selon la revendication 8 ou 9, caractérisé en ce qu'il comporte également des moyens (24-29) formant les histogrammes des valeurs de sortie de ladite unité de traitement spatial, ainsi que les histogrammes des inclinaisons de deux axes, à pente variable, de coordonnées dans un plan, des moyens pour repérer, dans chaque histogramme, un domaine de variation significative de la valeur traitée, afin de valider ce domaine en leur sortie et de déduire, pour l'ensemble des histogrammes, des signaux de sortie qui signalent et localisent une zone de la scène observée en mouvement relatif, si elle existe, ainsi que la vitesse et la direction orientée de ce mouvement si ladite zone se déplace effectivement par rapport à son environnement.

11. Dispositif selon la revendication 8, 9 ou 10, caractérisé en ce que lesdits moyens de lissage (15c, 15d) comportent une entrée qui reçoit ledit signal vidéo numérique S (PI) et calculent, pour les pixels successifs d'une trame de ce signal vidéo, un signal lissé (LO), dans lequel les variations temporelles du signal vidéo numérique d'entrée sont diminuées, par mise en oeuvre d'un signal de seuil (N) reçu sur une autre entrée et d'une constante de temps (CO) relative à chaque position de pixel d'une trame, dont la valeur est successivement modifiée afin que le lissage conserve, tout en la réduisant, la tendance de variation du signal vidéo numérique entrant, ces moyens de lissage coopérant avec l'unité de mémoire (16) qui reçoit, stocke et restitue les valeurs actualisées, pour chaque position de pixel d'une trame, du signal lissé et de ladite constante de temps et débitent, sur leurs sorties, au moins la succession, pour chaque position de pixel, des valeurs de la constante de temps actualisée et des valeurs d'un signal binaire de dépassement ou non-dépassement dudit seuil par la valeur absolue de la différence entre la valeur du pixel et sa valeur lissée.

12. Dispositif selon l'une quelconque des revendications 8 à 11, caractérisé en ce que ladite unité de traitement spatial (17, 18) réalisant la répartition matricielle, par lignes et par colonnes en nombre réduit, des sorties desdits moyens de lissage, à savoir des valeurs successives de la constante de temps (CO) et dudit signal binaire, comporte des moyens de repérage (17a) pour repérer, dans ladite répartition matricielle, une zone de pixels dans laquelle, à la fois, soit la valeur dudit signal binaire correspond à un dépassement de seuil et ladite constante de temps varie entre pixels voisins d'une valeur significative dans une direction, et pour produire des signaux de sortie indiquant la localisation de ladite zone et la vitesse et la direction orientée du déplacement dans ladite zone, soit la valeur dudit signal binaire correspond à un non-dépassement du seuil et ladite constante de temps ne varie pas entre pixels voisins.

13. Application du dispositif selon la revendication 10, 11 ou 12, pour aider la conduite d'un véhicule terrestre sur une route, respectivement d'un véhicule aérien au voisinage d'une piste d'aéroport, caractérisé en ce que le dispositif comporte en outre des moyens de représentation des bords droit Bd et gauche Bg de la route, respectivement de la piste, et des moyens pour orienter au moins un des axes, à pente variable, de coordonnées pour qu'il se maintienne sensiblement orthogonal au bord correspondant (position Po).

14. Application du dispositif selon l'une quelconque des revendications 1 à 12, pour surveiller l'état de vigilance du conducteur d'un véhicule automobile, afin de détecter une tendance éventuelle à la somnolence de celui-ci, consistant à:

- générer un signal vidéo numérique représentant
  - initialement les images successives de la face du conducteur et

- ensuite, d'une manière continue et en temps réel, les images successives de seulement les yeux du conducteur;

- traiter ledit signal vidéo relatif seulement aux yeux, du conducteur afin de, successivement et en temps réel,

- détecter, dans lesdites images de seulement les yeux, les mouvements verticaux des paupières représentant le clignement de celles-ci,
- déterminer les cadences successives de ces mouvements et
- détecter les cadences qui sont inférieures à un seuil de clignement des paupières qui correspond à la transition entre l'état éveillé et l'état de somnolence du conducteur, et

- déclencher un signal d'alarme apte à alerter le conducteur dès que lesdites cadences franchissent ledit seuil.

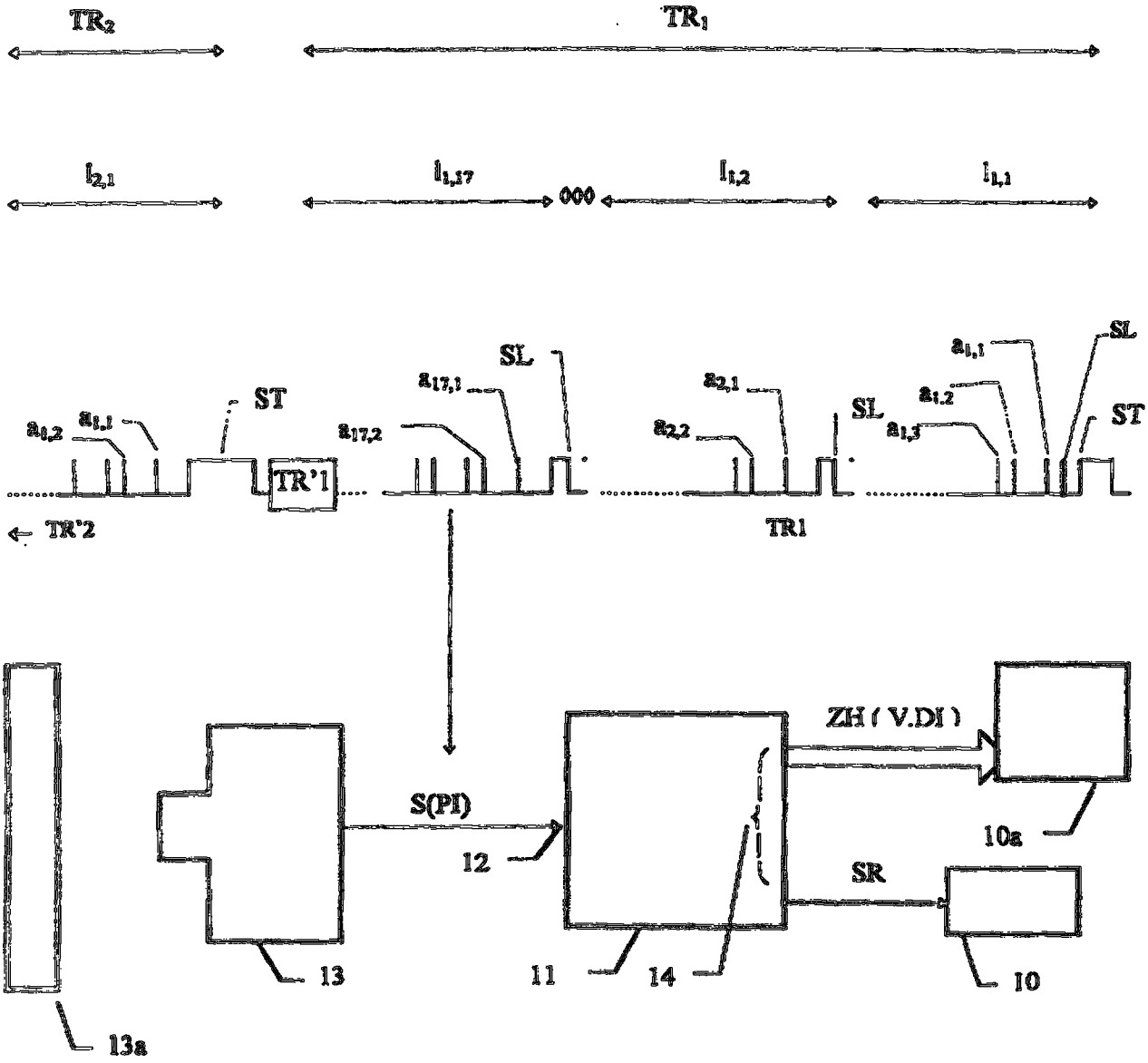

TR2
TR1
l2,1
l1,17
000
l1,2
l1,1
a1,2
a1,1
ST
a17,1
SL
a17,2
TR'1
TR'2
a2,1
a2,2
SL
a1,1
a1,2
a1,3
SL
ST
TR1
S(PI)
12
13
13a
11
14
ZH (V.DI)
10a
SR
10

FIG. 1

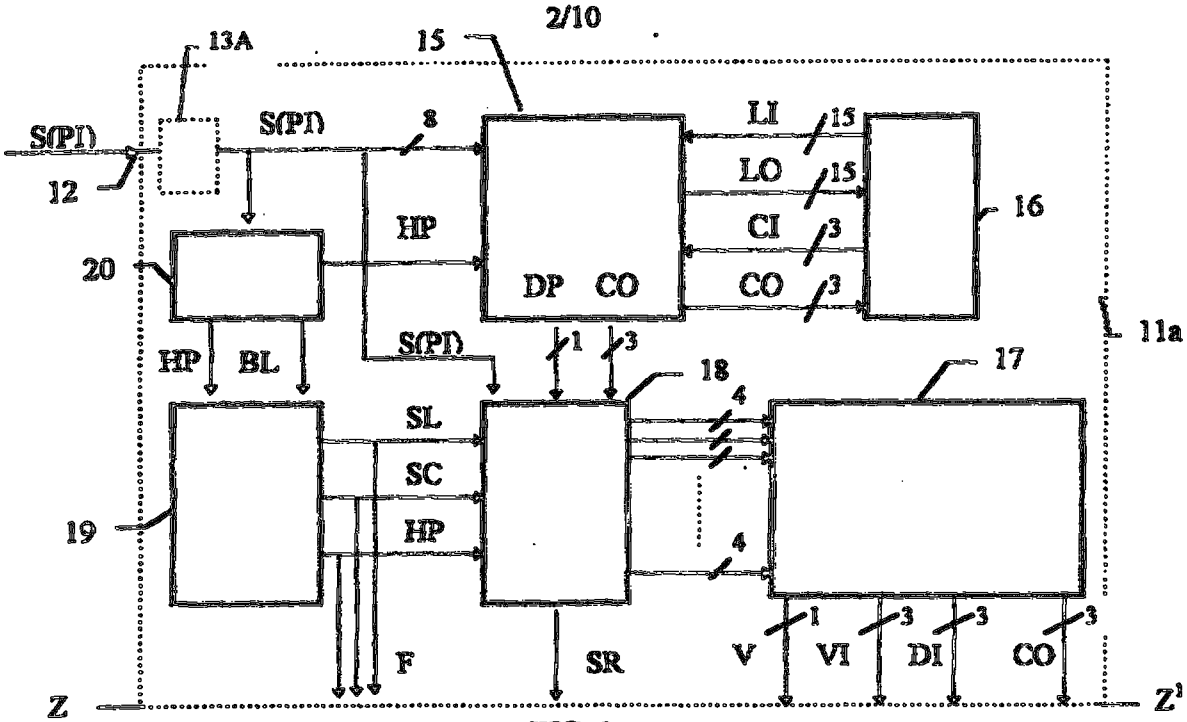
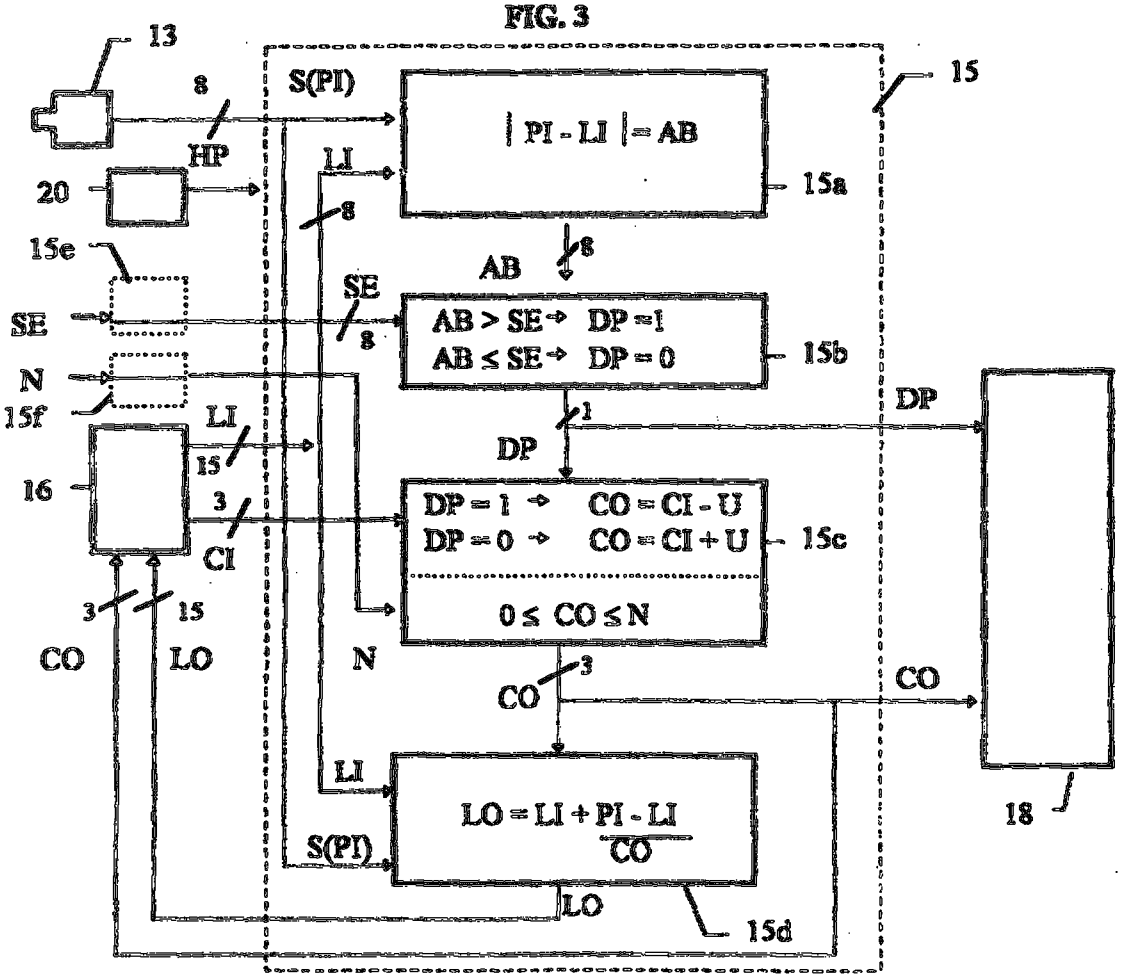

13A
15
S(PI)
S(PI)
8
LI
15
LO
15
CI
3
CO
3
16
12
HP
20
DP CO
11a
HP
BL
S(PI)
1
3
18
17
4
SL
SC
HP
19
4
F
SR
V
1
VI
3
DI
3
CO
3
Z
Z'

FIG. 2

FIG. 3


13
8
S(PI)
| PI - LI | = AB
15
15a
HP
LI
20
8
AB
8
15e
SE
SE
AB > SE → DP = 1
AB ≤ SE → DP = 0
15b
8
N
15f
LI
DP
1
DP
15
16
DP = 1 → CO = CI - U
DP = 0 → CO = CI + U
15c
3
CI
0 ≤ CO ≤ N
3
15
CO
LO
N
CO
3
CO
LI
LO = LI + PI - LI
CO
18
S(PI)
LO
15d

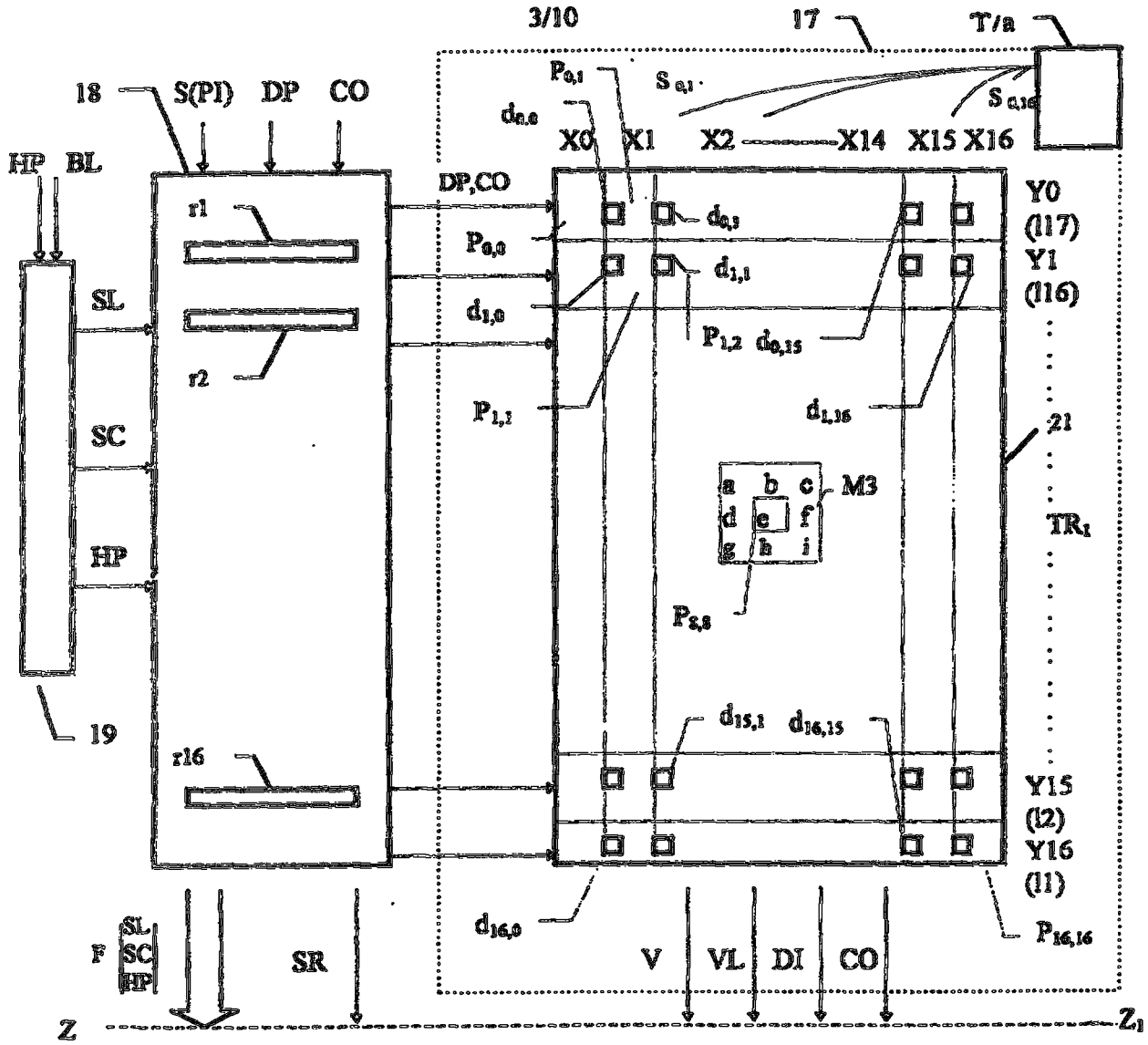


FIG. 4

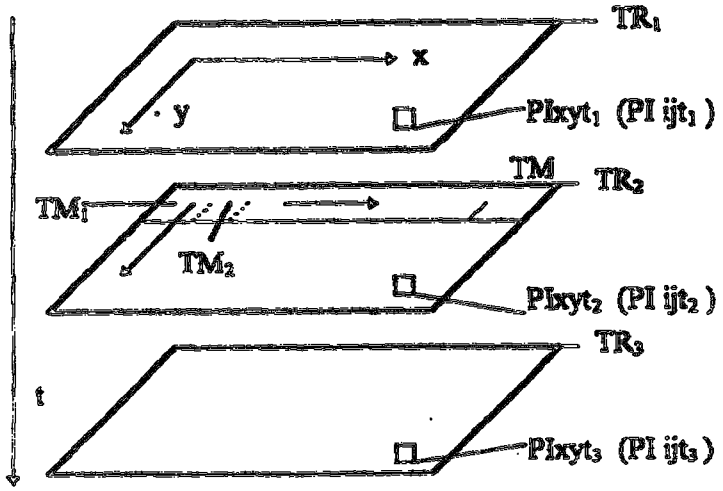


FIG. 5

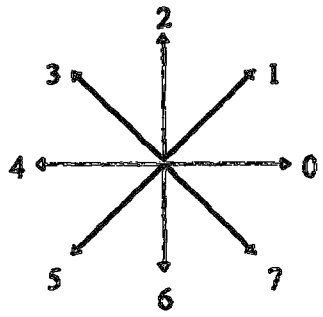
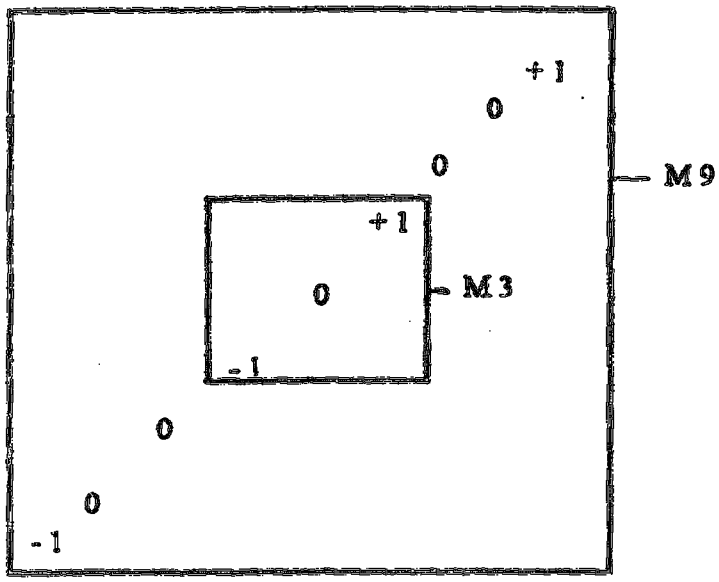


FIG. 6

FIG. 7



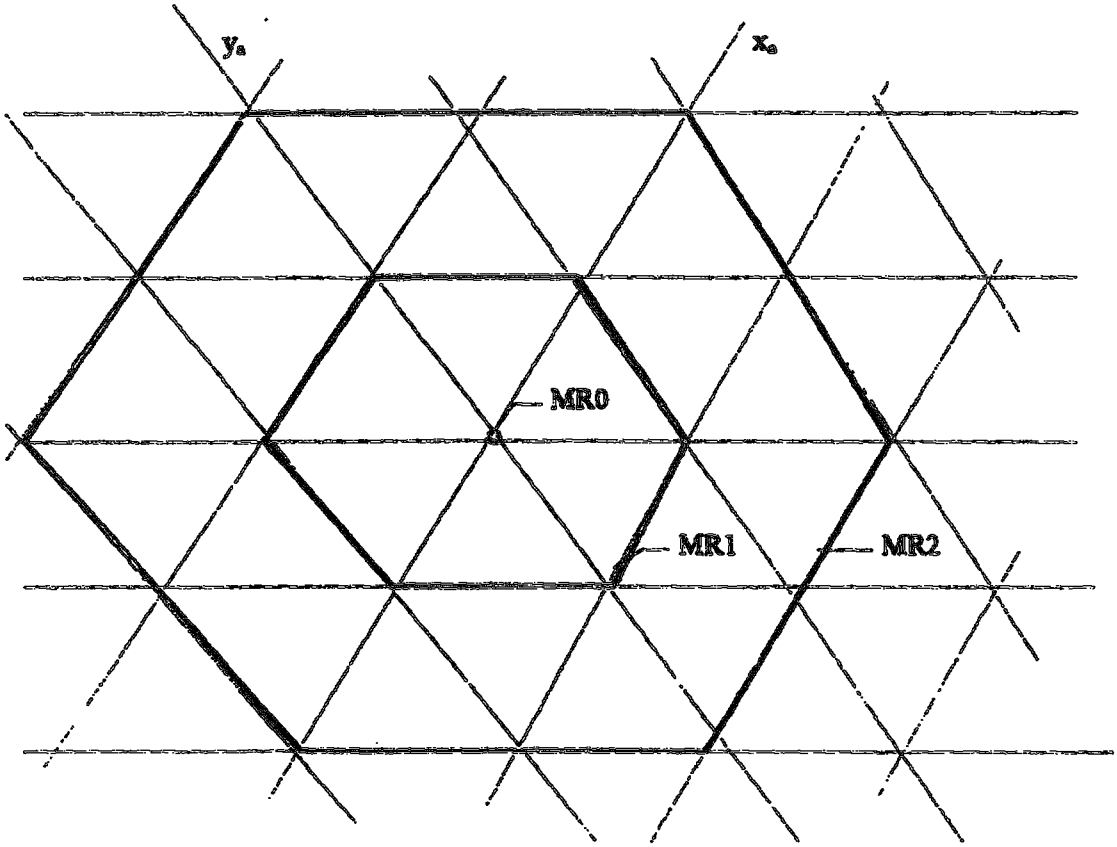


FIG. 8

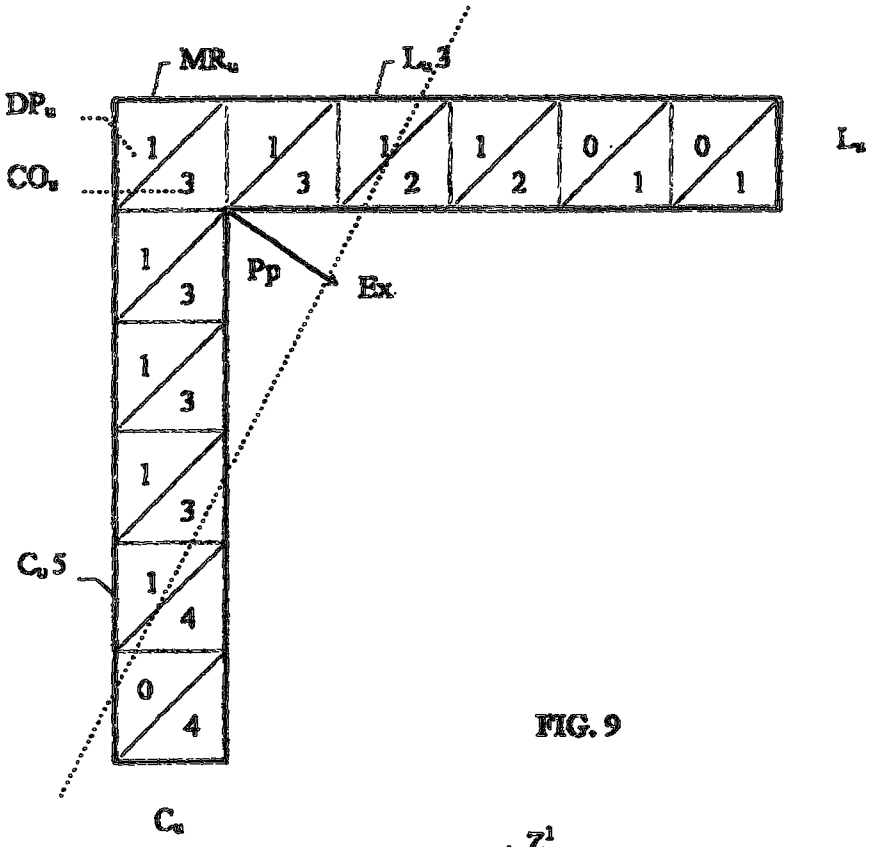

MR
L 3
DP
CO
1 3
1 3
1 2
1 2
0 1
0 1
L
Pp
Ex
1 3
1 3
1 3
C 5
1 4
0 4
C

FIG. 9

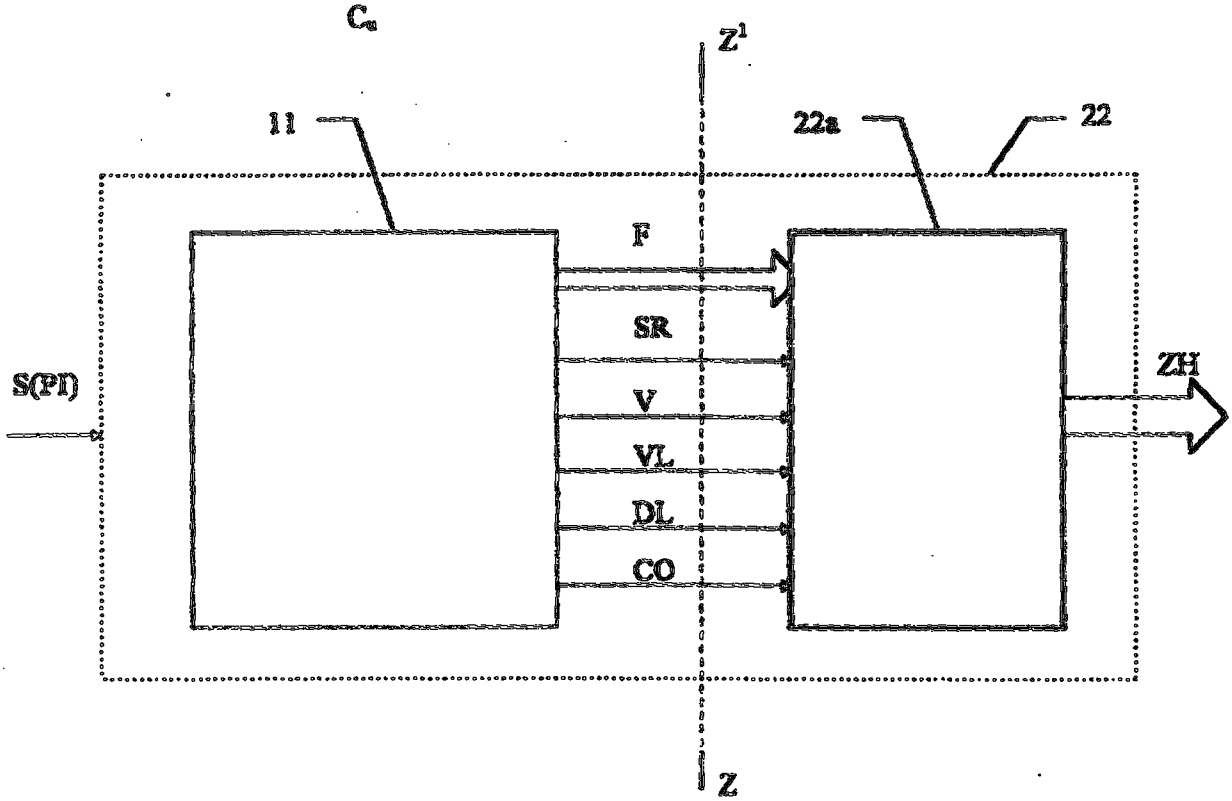

Z'
11
22a
22
F
SR
V
VL
DL
CO
S(PI)
ZH
Z

FIG. 10

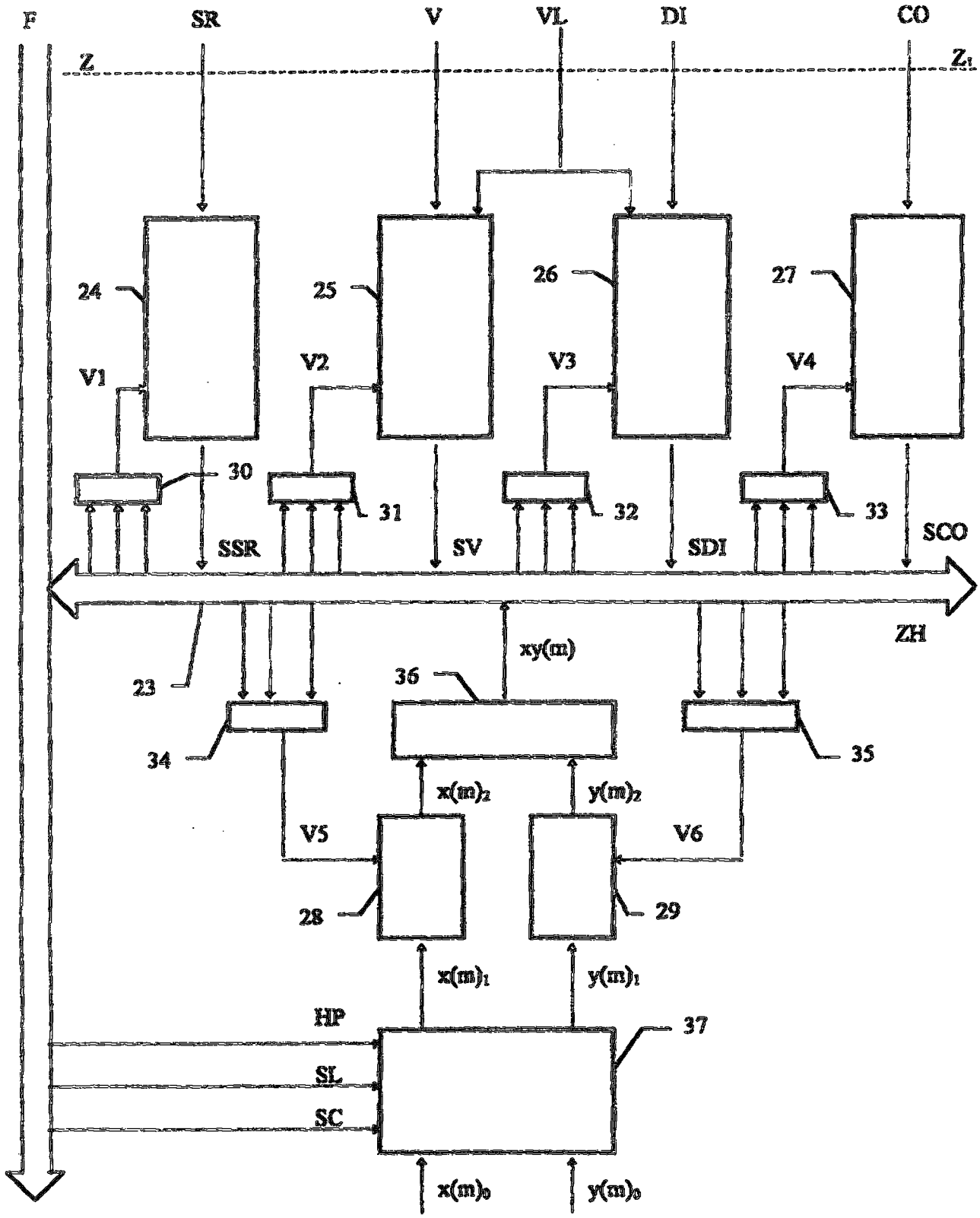

F
SR
V
VL
DI
CO
Z
Z1
24
25
26
27
V1
V2
V3
V4
30
31
32
33
SSR
SV
SDI
SCO
ZH
xy(m)
23
36
34
35
x(m)2
y(m)2
V5
V6
28
29
x(m)1
y(m)1
HP
37
SL
SC
x(m)0
y(m)0

FIG. 11

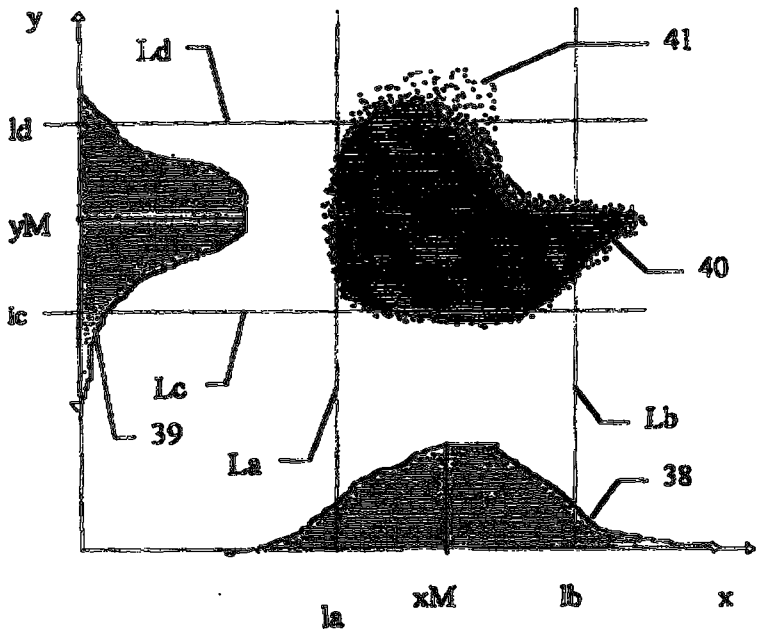
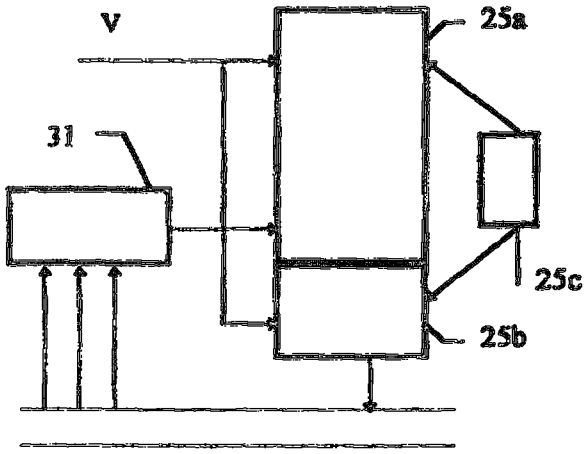


FIG. 12

FIG. 13



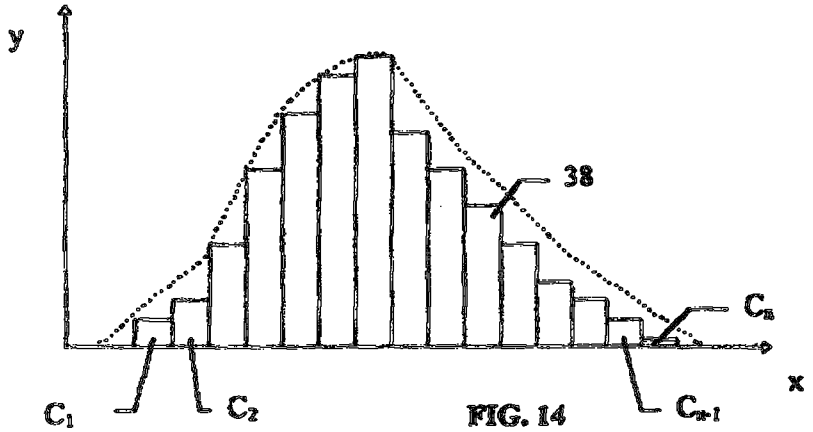
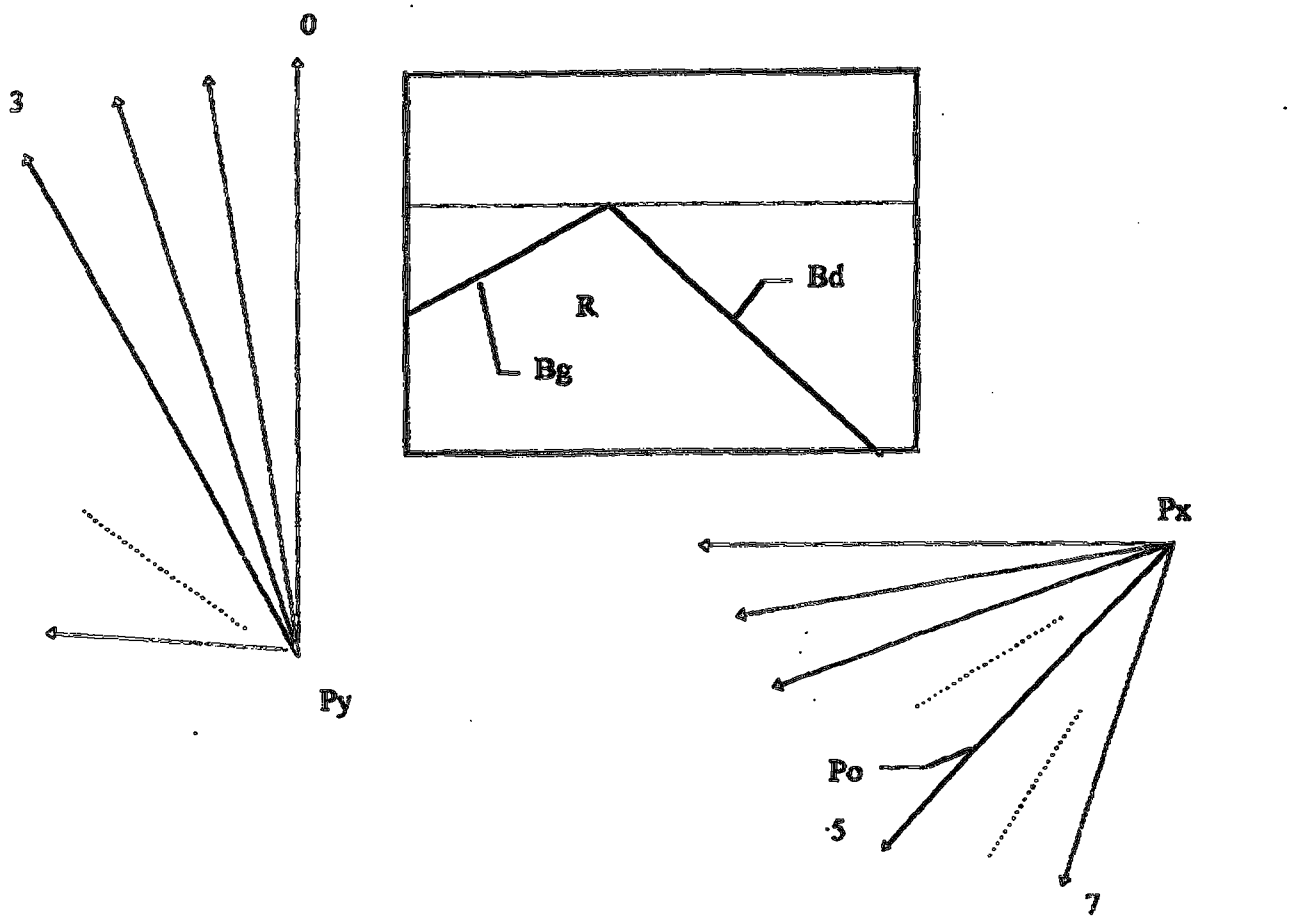


FIG. 14

FIG. 15



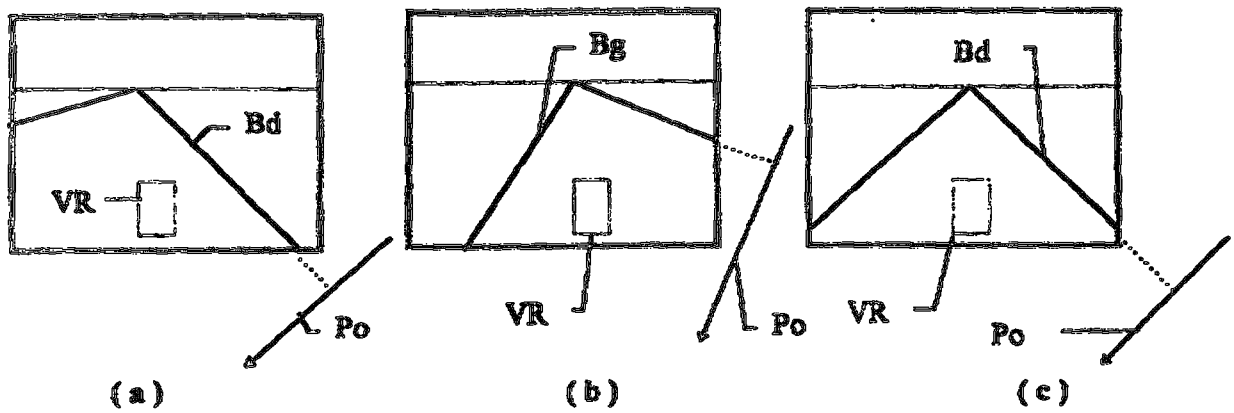
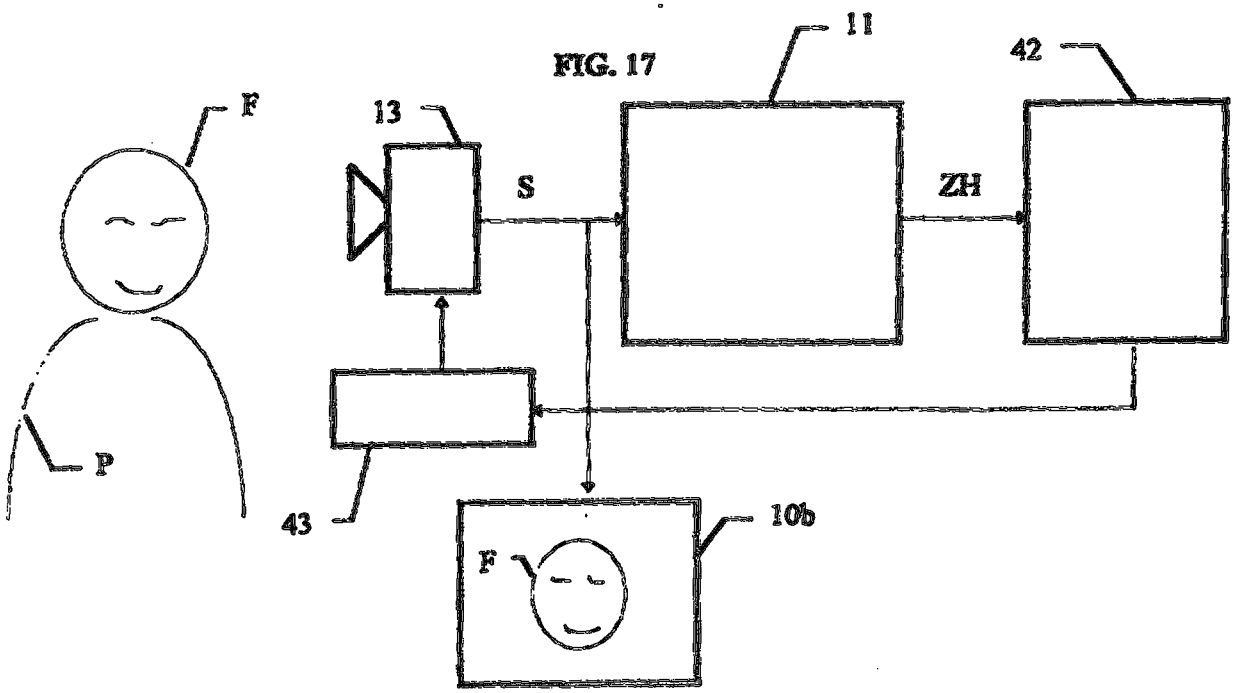


FIG. 16

FIG. 17



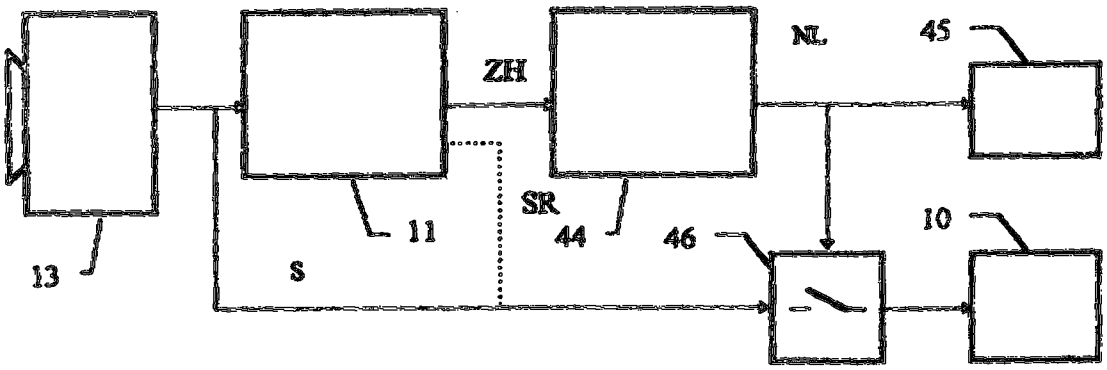


FIG. 18

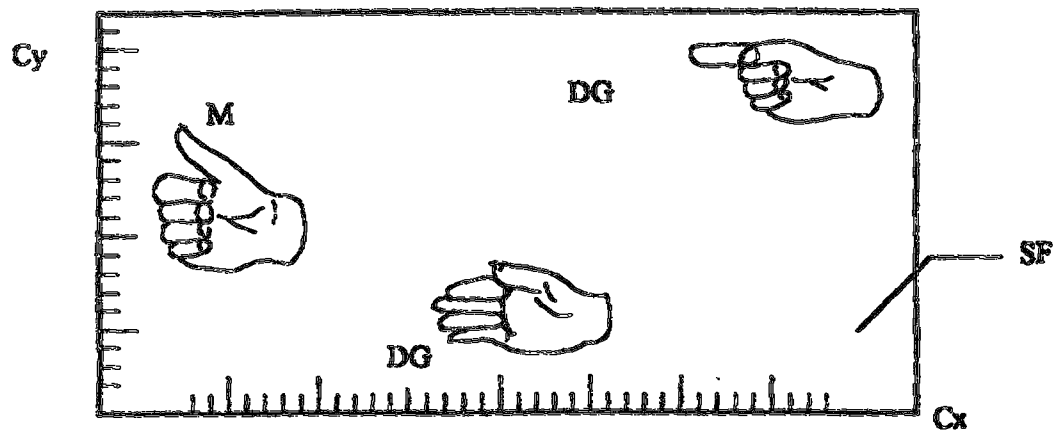


FIG. 19

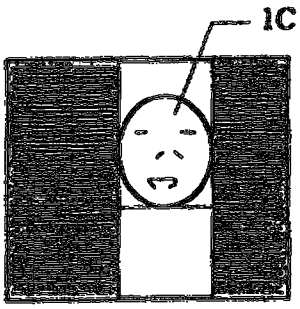


FIG.20



FIG.21

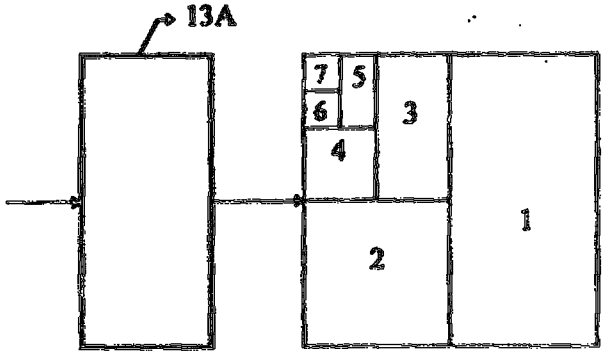


FIG.22

PROCESS AND DEVICE OPERATING IN REAL TIME FOR IDENTIFYING AND LOCALIZING AN AREA IN RELATIVE MOVEMENT IN A SCENE, AND FOR DETERMINING THE SPEED AND DIRECTION OF THE DISPLACEMENT

The purpose of this invention is a process and device for identifying and localizing an area in relative movement in a scene and for determining the speed and oriented direction of this relative movement, in real time.

Relative movement means a movement of said area (which may be composed of an "object" in the broadest sense of the term including a living being or a portion of a living being, for example a hand) in an approximately motionless environment, or more or less complete immobility of said area (or "object") in an environment that is at least partially in movement.

The invention relates to the processing of a digital video output signal from an observation system composed of an optical imput system or objective capable of forming an image of the observed scene, and by an optoelectronic conversion system or sensor capable of converting said image that it receives into a digital output signal.

In general, the observation system is composed of a video camera or camescope that observes the scene to be monitored (said digital output signal then being composed of the digital video signal output by a camera with a digital output or of the output of an analog/digital converter, the input of which is connected to the output of a camera outputting an analog video signal.

The observation system may also be composed of the objective of an optical instrument (binoculars, sighting refracting instrument, view-finder), from which at least a portion of the outgoing light beam is sampled and by a photoelectronic sensor, for example of the CCD

or CMOS type, with the normal associated electronic system, sensor receiving the image formed by said portion of the light beam and converting it through the associated electronic system into a digital video output signal.

The invention consists mainly of processing the digital video output signal from an observation system, particularly a video camera, with a digital output to determine signals signaling the existence and localization of an area in relative movement in said scene, and the speed and oriented direction of the displacement if said area is actually moving in said scene relatively to an approximately motionless environment, in real time.

The best system yet developed for identifying and localizing an object in relative movement and for determining its speed and oriented direction of displacement is animal or human sight, for example a hunter in his hide localizing the displacement of an animal, and the direction and speed of this displacement.

Prior art surveillance devices like artificial retinas, either analog (Giocomo Indiveri et al, in Proceedings of MicroNeuro'96 pp 15 to 22), or digital (Pierre-François Rüedi in Proceedings of MicroNeuro'96 pp 23 to 29), have been proposed, but the first article deals with detectors and analog units with complex structures, and the second article deals with means of identifying the edges of an object; furthermore, very fast and high capacity memories are used in the devices described in order to operate in real time, and limited information is obtained about moving areas or objects.

Thus the suggestion was made of memorizing a frame output signal from a video camera, or similar instrument, composed of a sequence of data relative to pixels representative of the scene observed by the camera at time t_0 in a first two-dimensional memory, and then

the video signal for the next corresponding frame representing said scene at time t_1, in a second two-dimensional memory. If an object has moved between t_0 and t_1, on the one hand, the distance d by which the object has moved in the scene between t_1 and t_0 and, on the other hand, the time $T = t_1 - t_0$ between the beginnings of two successive corresponding frames relative to the same pixels are determined. The displacement speed is then equal to d/T. This type of system requires a very large total memory capacity if it is required to obtain precise speed and oriented direction indications characterizing the displacement. Furthermore, there is some delay in obtaining the speed and displacement direction indications; this information is not available until time $t_1 + R$, where R is the time necessary for the calculations for the period $t_0 - t_1$. These two disadvantages (the need for a large memory capacity and the delay in obtaining the required information) limit applications of this type of system.

Furthermore, French patent No. 2.611.063, of which one of the inventors (Mr. Patrick Pirim) is the inventor of the present invention, describes a process and a device for real time processing of a sequenced data flow, composed in particular of the output signal from a camescope, in order to perform data compression. According to this prior patent, the histogram of signal levels is formed using a classification law for a first sequence, the representative Gauss function associated with this histogram is memorized and the maximum and minimum levels are extracted from it, the levels of the next sequence or second sequence are compared with said signal levels for the first sequence memorized with a constant time constant identical for each pixel, a binary classification signal is generated that characterizes the

next said sequence with reference to the classification law, an auxiliary signal is generated from this binary signal representative of the duration and position of a range of significant values, and finally said auxiliary signal is used to generate a signal localizing the range with the longest duration, called the dominant range, and these operations are repeated for subsequent sequences of the sequenced signal. This process and this device of classification enable data compression, keeping only interesting parameters in the processed flow of sequenced data. In particular, this process is capable of processing a digital video signal representative of a video image in order to extract and localize at least one characteristic of at least one area in said image. It is thus possible to classify brightness and/or chrominance levels of the signal and to characterize and localize an object in the image.

United States patent No. 5:488.430 detects and estimates a displacement by determining horizontal and vertical changes in the image of the observed area separately. Difference signals are used to detect movements from right to left or from left to right, or from top to bottom or bottom to top, in the horizontal and vertical directions respectively, firstly by carrying out the EXCLUSIVE OR logical function on horizontal/vertical difference signals and frame difference signals, and secondly by using a ratio of the sums of horizontal/vertical signals and the sums of frame difference signals with respect to a K x 3 window. In this US patent No. 5.488.430, calculated values of the image along two orthogonal horizontal and vertical directions are used with an identical repetitive difference K in these two orthogonal directions, this difference K being defined as a function of the displacement speeds that are to be determined. The device according to this US patent determines the

direction of movements along each of the two orthogonal directions by applying a set of calculation operations indicated in columns 12 (at the beginning and end) and 13 (at the beginning) to the difference signals, which requires very complex (and therefore difficult to execute) electronic comPuting units performing particularly division, multiplication and summation ; furthermore, additional complex computing units are necessary to obtain the speed and oriented direction of displacement (extraction of square root to obtain the amplitude of the speed and calculation of the arctan function to obtain the oriented direction), starting from projections on the two horizontal and vertical axes. Finally patent No. 5 488 430 does not include any means of smoothing pixel values by means of a time constant, variable for each pixel, in order to compensate for excessively fast variations of these values.

On the contrary, the process according to the invention is embodied by using a digital type of device, object of the invention, which has a relatively simple structure and a relatively small memory capacity, and by which the required information can be obtained quickly with a wide variety of results and applications (corresponding to a complete half-image for interlaced frames or a complete image, depending on the application).

An article by Alberto Tomita Sr. and Rokuya Ishii, entitled "Hand Shape Extraction from a Sequence of Digitized Gray-Scale Images" in Institute of Electrical and Electronics Engineers, vol. 3, 1994, p 1925-1930, implements a movement detection by subtraction between successive images followed by the formation of histograms based on the shape of a human hand in order to extract the shape of a human hand in a digitized scene. The histogram analysis is based on a gray scale inherent to the human hand. It does not include any means of forming

histograms of the plane coordinates. The sole purpose of the authors of this article is to detect the displacements of a human hand, for example in order to replace the normal mouse by a hand, the displacements of which are identified, for inputting data into a computer.

On the contrary, the present invention is not limited to detecting the displacement of a hand, but is capable of detecting the relative displacement of any object (in the widest sense of the term) in a scene, and does not use histograms based on the gray values of a hand, but rather the histograms of some particular digital variables representative of the displacement (if any) and histograms of plane coordinates.

According to this invention:

- a digital video output signal from an observation system is processed, this signal being composed of a sequence of frames (corresponding to a half-image in the case of two interlaced frames per image or to a complete image in the case of a single frame per image) in the known manner, each consisting of a given number of successive lines, and a given number of pixels or image points in each of these lines,
- in order to obtain, using a relatively small memory capacity, signals capable of indicating if there is an area in relative movement in the observed scene and, if so, to specify the localization, and speed and oriented direction of said area if it is actually moving with respect to its environment,
- by generating two digital signals, one characteristic of the significant variation or non-variation of the pixel signal, for the same pixel location, between two successive corresponding frames, and the other characteristic of the amplitude of this variation, if any, and performing a matrix distribution of these two signals for the pixels in a portion of a frame at a same instant.

It is a first object of the invention to provide a process operating in real time for identifying and localizing an area in relative movement in a scene observed by an observation system with an output constituted by a digital video signal of the type comprising a succession of corresponding frames, each composed of a succession of lines, each composed of a succession of pixels, and for determining the speed and oriented direction of the displacement, said process being characterized in that it consists in carrying out the following sequence of operations on the digital video output signal:

- a smoothing processing on said digital video output signal using a digital time constant, the numerical value of which may be modified independently for each pixel of said output signal,
- a memorization, on the one hand, of a frame of said output signal after smoothing and, on the other hand, of the smoothing time constant associated with said frame,
- a temporal processing for each pixel position, consisting in determining firstly the existence, and secondly the amplitude of a significant variation in the amplitude of the pixel signal between the current frame and the immediately preceding smoothed and memorized frame, and generating two digital signals, the first signal being a binary, or single bit, signal with two possible values, one of which represents the presence and the other represents the absence of a significant variation between two successive frames, the value of said binary signal modifying the memorized value of said time constant in order to reduce it if said signal represents a significant variation and in order to increase it if said signal does not represent such a variation, the reduction or increase being done in a quantified manner, whereas the second digital

signal, i.e. the amplitude signal, is a multibit signal with a limited number of bits quantifying the amplitude of this variation; and

- a spatial processing consisting of the following steps for each frame of the digital video input signal :
  - ∘ distributing only the values of a fraction of the pixels in a frame at a given observation instant (a fraction which is scanned across said matrix throughout the duration of the frame) firstly of said binary signal and secondly of said digital amplitude signal, into a matrix with a number of rows and columns that is small compared with the number of lines and the number of pixels per line in the video signal, respectively, in order to characterize the values of the pixels,
  - ∘ determining in this double instantaneous matrix representation, a particular area in which said binary signal has the value sought for, representing the presence or absence of a significant variation, and said digital amplitude signal varies, or does not vary, by a significant value for adjacent pixels in the matrix along an oriented direction starting from an origin pixel, within the same portion of the frame and therefore at the same observation instant, and
  - ∘ generating signals representing the existence and localization of the area in relative displacement and the relative inter-frame speed and the oriented direction of this displacement, if any, relative to its environment, inferred from the instantaneous matrix distribution of these two binary and amplitude digital signals.

Preferably, the process according to the invention is characterized in that it also consists of:

- ∘ forming firstly histograms of the signal values distributed in matrices, and secondly histograms of

the inclinations of two coordinate axes, with variable slope, in a plane,

- identifying a domain in which there is a significant variation of the processed values in each histogram formed, and
- for each identified domain, deducing whether or not there is an area in relative movement, and if so its localization, speed and oriented direction.

In specific embodiments:

- said matrix is a square matrix with the same odd number of rows and columns $(2l + 1)$, and nested matrices containing 3 x 3, 5 x 5, 7 x 7, ...$(2l + 1)$ x $(2l + 1)$ elements, centered on the center of this square matrix, are considered in order to determine the smallest nested matrix in which said digital signal varies along an oriented direction starting from said center, the value of said binary signal showing that the threshold limit along this direction has been exceeded,
- said matrix is a hexagonal matrix, and nested hexagonal matrices of increasing size, centered on the center of this hexagonal matrix, are considered, in order to determine the smallest nested matrix in which said digital signal varies along an oriented direction,
- said matrix is an inverted L matrix with a single row and a single column, and nested matrices containing 3 x 3 pixels, 5 x 5 pixels, 7 x 7 pixels,... $(2l + 1)$ x $(2l + 1)$ pixels are considered for the single row and the single column, in order to determine the smallest matrix in which the signal varies along an oriented direction, namely the line with the steepest slope and constant quantification.

Advantageously, said time constant is of the form 2^p, where p is a number less than 16, and which can then be expressed in not more than 4 bits, the time constant

being reduced or increased by subtracting or adding one unit to *p*.

In this case, if it is required, successive decreasing portions of complete frames are considered using the Mallat time-scale algorithm and the largest of these portions, which gives displacement, speed and orientation indications compatible with the value of *p*, is selected.

Another object of the invention is a device operating in real time for identifying and localizing an area in relative movement in a scene observed by an observation system with an output constituted by a digital video signal of the type comprising a succession of corresponding frames, successive lines in each corresponding frame and successive pixels in each line, and for determining the speed and oriented direction of the displacement, by using the process described above, said device receiving said video output signal as input and being characterized in that it comprises in combination:

- means for smoothing said digital video output signal making use of a digital time constant, the numerical value of which may be modified independently for each of said output signal pixels;
- means for memorizing firstly a frame of said output signal after smoothing and secondly the smoothing time constant associated with said frame,
- a temporal processing unit to analyze the variations with time of the amplitude of the pixel signal, for the same pixel position, between the current frame and the immediately preceding smoothed and memorized frame of said digital signal, said unit comprising a memory capable of receiving, storing and retrieving information about the corresponding previous smoothed frame, together with means of comparison to determine if the absolute value of the difference, between the

current pixel signal and a representative value of the pixel signal for the same pixel position in the previous frame stored in said memory, exceeds or does not exceed a threshold, for each position of the pixel in the input video signal frame, by generating a binary, or single bit, signal with two values, one of these two values showing that the threshold is exceeded and the other showing that it is not exceeded, and computing means capable of determining a multibit digital amplitude signal with a small number of bits, the value of which depends on the amplitude of the variation in the value of the same pixel between the current frame and the immediately preceding, smoothed and memorized, frame of the digital video signal; and

- a spatial processing unit, into which the outputs from said successive binary and digital amplitude signals for the pixels in a given frame produced by the temporal processing unit are input, said spatial processing unit comprising means for characterizing the amplitude values of the pixels, these means distributing only said binary and digital amplitude signals relative to the same instant, i.e. to a single frame, into a matrix with fewer rows and columns than the number of lines and the number of pixels per line respectively in a frame of said digital video signal, the frame being scanned across said matrix throughout the duration of a frame, spotting means for determining an area of pixels within said matrix in which, at this instant, the binary signal has the searched value, and means of determining an area of pixels in said matrix, into which, at the same instant, the digital amplitude signal for adjacent pixels varies by a significant quantity, and means which, in response to the indications from the two preceding means, generate signals representing this area of pixels and therefore the existence and localization of an area in relative

movement within the observed scene and the relative inter-frame speed and oriented direction of this area when it actually moves relative to its environment.

Preferably, the spatial processing unit will also include an output delivering a delayed digital video signal consisting of the input digital video signal with a delay equal to the duration of the rows of a matrix minus the duration of a row, in order to supply an output signal contemporary with the analysis of the matrix in said temporal processing unit.

Preferably, the spatial processing unit of the device according to the invention, includes delay means in cascade, each of which imposes a delay equal to the time difference between the beginning of two successive lines, and delay means in cascade for each line, each importing a delay equal to the time difference between two successive pixels in a line, the outputs from all registers and the input to the first registers on each line delivering at a given instant the values of said binary signal and said digital amplitude signal at the same instant, to said identification means.

Advantageously, the device according to the invention also comprises means for forming histograms of the output values from said spatial processing unit, and histograms of the inclinations of two, variable slope, coordinate axes within a plane, means for identifying a domain, in each histogram, in which there is a significant variation in the processed value, in order to validate this domain at their output, and to deduce, for all histograms, output signals that identify and localize an area of relative movement in the observed scene, if any, and the speed and oriented direction of this movement if said area is effectively moving with respect to its environment.

If it is required to detect the movement of an object in an approximately motionless environment, it is

necessary to determine the part of the matrix in which the value of the binary signal shows that the limit has been exceeded and at the same time the digital amplitude signal varies by a significant value between adjacent pixels in a frame.

On the contrary, if it is required to detect a motionless object in an environment that is effectively in motion, it is necessary to determine the part of the matrix in which the value of the binary signal shows that the limit has not been exceeded and at the same time the digital amplitude signal does not vary between adjacent pixels in a frame.

Preferably in the device for identifying, localizing and determining the speed and oriented direction of displacement of an area in relative movement in a scene, embodying the above mentioned process:

- said smoothing means include an input receiving said digital video signal and derive, for each successive pixel in a frame of said video signal, a smoothed signal in which the temporal variations of the input digital video signal are reduced by using a threshold signal received on another input and a time constant related to each pixel position in a frame, the value of which is modified successively so that the smoothing maintains, although reduced, the trend in the variations of the incoming digital video signal, said smoothing means working in cooperation with said memory unit that receives, stores and retrieves the updated values of the smoothed signal and of said time constant, for each portion of pixels in a frame, and output, for each pixel position, at least the succession of values of the updated time constant and of the values of the binary signal that shows whether or not said threshold is exceeded by the absolute value of the difference between the value of the pixel and its smoothed value; said spatial processing assembly

carrying out the matrix distribution along lines and columns in a reduced number, of the outputs from said smoothing means, namely the successive values of the time constant and of said binary signal ;

- means are provided for identifying, in said matrix distribution, an area of pixels in which simultaneously:

either the value of said binary signal corresponds to a threshold being exceeded and said time constant varies between adjacent pixels by a significant value in one direction and for producing output signals indicating the localization of said area and the speed and oriented direction of displacement in this area ;

- or the value of said binary signal corresponds to a limit not being exceeded and said time constant does not vary between adjacent pixels and for producing output signals indicating the localization of said area.

In the preferred embodiments:

- said smoothing means include, in combination with a video memory or field memory that stores successive values, for each pixel in the frame, of said time constant and the smoothed digital video signal, means for calculating, for each pixel, the absolute value of the difference between the value of the digital video signal output from the camera and the value of the previously smoothed digital video signal, means for comparing said difference with a threshold, and for generating a binary signal, one of the two values of which indicates that said limit has been exceeded and the other value indicates that it has not been exceeded; means for updating the time constant receiving the immediately preceding value of the time constant from said memory and reducing it if it receives a binary signal with a value showing that the limit has been exceeded, or increase it if the value of

the binary signal indicates that it has not been exceeded, the reduction or increase being however not performed if it would lead respectively to a negative value or to a value exceeding a threshold value, and means for updating the smoothed value of the digital video signal that algebraically add, to the prior value of this smoothed signal received from said memory, the quotient of the difference between the digital video signal output from the camera and the previous smoothed digital video signal from said memory by a factor equal to value of the previous time constant retrieved by from said memory;

- said time constant is in the form 2^p, where p is an integer less than 16, and which can therefore be represented by not more than 4 bits, the reduction or increase in the time constant being done by subtracting or adding one unit to p;
- said matrix distribution means include in combination with delay means applying successive delays, equal to the duration of one line of the video signal, to the digital video signal originating from the camera so as to deliver on successive outputs this video signal delayed by one delay period, two delay periods and so on, up to a number of delay periods equal to the number of lines in the spatial distribution matrix minus one, means for performing a matrix distribution along successive rows of the matrix receiving firstly, the digital video signal from the camera without a delay and this progressively delayed signal output from the delaying means outputs and secondly, values of the time constant and said binary signal received from said smoothing means in order for producing a matrix distribution, in rows and columns, of said values, at a given instant, of the time constant and binary signal for pixels in a portion of the frame of the digital video signal with the same size as the matrix;

- the matrix distribution means comprise a series of digital signal conductors, i.e. one for each row in the distribution matrix, each with series-connected shift registers, each imposing a delay equal to the time difference between two successive pixels in a line of the digital video signal, the position of a pixel distributed in the matrix being determined by a point in the matrix located before a delay register, the number per line thereof being equal to the number of columns in the matrix minus one, and by a point downstream the furthest downstream register;
- said means of detecting a moving area in said matrix, by detecting the simultaneous presence of a value of the binary signal indicating that a limit was exceeded and a variation in the value of the time constant, include means for determining, along discrete digitized oriented directions, the slope of the variation in the value of the time constant close to a pixel at the center of said matrix, forming the origin for said directions, and means for selecting the steepest slope of this variation close to this center-origin and determining the oriented direction, taking account of a selection criterion for selecting the direction if there is more than one direction with same maximum slope of this variation, these latter means outputting signals representative of the speed and oriented direction of the displacement in the moving area, with a validation signal indicating that these speed and direction signals are validated, together with the value of the time constant.

Preferably, if included in the device according to the invention, the means for creating the histograms each include:

- inputs receiving the signal from which the histogram is formed and a validation signal originating from the matrix distribution means; and

- means for establishing two linear unidimensional histograms for the two plane coordinates and for combining these two linear histograms into a surface histogram representing the area in which there is a significant variation in the input signal, and
- an output delivering a signal representative of said area.

Furthermore, the means for forming the histograms preferably include:

- means for calculating a reference change, the inputs of which receive a line sequence signal, a column sequence signal and a pixel clock signal and the output of which is representative of the reference change ;
- two means for forming histograms for two axes, receiving the two reference signals and forming histograms of these axes; and
- area means into which the outputs from the two histogram formation means for the two axes are input, and which output a signal containing global slope information for the two axes.

In some applications, said time constant may be composed of the order number of the intervals, in limited number and with a gradually increasing magnitude, into which the absolute value of the difference between the value of the current pixel and the immediately preceding value of the same pixel after smoothing is divided, for each pixel position.

Advantageously, for determining the binary signal of overshooting, or not overshooting, of the threshold the absolute value of the difference between the value of the current input pixel and the immediately preceding smoothed value of the same pixel originating from the memory in compared with a threshold value, for each pixel position.

As concerns the digital amplitude signal, it is preferably generated as an integer representing the

tendency of the value of the current pixel to approach the immediately preceding smoothed value of this pixel, for each pixel position.

One of the characteristics of the invention consists of the fact that, for determining both the binary signal and the digital amplitude signal, use is made, as value of the immediately preceeding pixel, of a stored smoothed value thereof, in order to reduce any excessive variations of this pixel signal with time that may exist in the digital video input signal originating from the video camera or any other observation device with a digital output.

It is known that the effect of a smoothing operation is to gradually substitute, for a digital signal with large amplitude variations with time, a signal with smaller variations and therefore more easily quantifiable with a small number of steps, and therefore bits, in the smoothed digital amplitude signal.

We will now describe a preferred embodiment, given for illustrative purposes and in no way limitative, of a device according to the invention embodying the process according to the invention, with reference to the appended drawings.

On the drawings :

Figure 1 illustrates very diagrammatically the entire system according to the invention with its inputs and outputs, and the input signal for this system.

Figure 2 represents, with functional blocks, the main units in a device according to the invention forming a temporal and spatial processing assembly.

Figures 3 and 4 illustrate functional diagrams of the temporal calculation processing assembly and the spatial matrix distribution processing assembly respectively, forming part of the device of figure 2.

Figure 5 shows diagrammatically temporal processing and spatial processing in a system according to the invention.

Figure 6 illustrates the numerical values, according to the Freeman code, of the discrete directions starting from an origin center in the matrix of figure 4.

Figure 7 illustrates two small nested matrices inside the temporal distribution matrix.

Figures 8 and 9 describe two other types of matrix, namely hexagonal and reverse L matrices respectively.

Figure 10 shows diagrammatically the assembly of the device in figure 2, along line Z - Z^1, with an additional assembly, according to a preferred embodiment of the invention.

Figure 11 shows said additional assembly in figure 10, in block form; figures 2 and figure 11 being assembled along line Z - Z^1 in dot and dashed lines shown on figures 2 and 10.

Figure 12 shows the formation of two single dimensional histograms, and, starting from these histograms, a two-dimensional histogram of a moving area, for an input signal.

Figure 13 shows more details of a histogram formation and processing block or unit, and its associated linear combination block.

Figure 14 illustrates a single dimensional histogram.

Figures 15 and 16 illustrate the use of a variation in the observation slope of a scene.

Figures 17, 18 and 19 illustrate other possible applications of the device according to the invention, in block form for figures 17 and 18, figure 17 concerning video conference, figure 18 the monitoring of highways, motorways or main roads and figure 19 an operator controlling a machine by hand movements.

Figures 20 and 21 diagrammatically represent an application of the invention to the monitoring of a car driver falling asleep.

Figure 22 shows the transformed image according to the Mallat diagram.

Referring firstly to figure 1, it can be seen that, in the preferred embodiment, device 11 according to the invention comprises firstly an input 12 that receives a digital video signal S originating from a video camera or camescope 13, with one or several CMOS type CCD sensors, having a direct digital output, or an analog output converted by an analog/digital converter into a digital output. This signal S is composed, in a known manner, by a succession of pairs of interlaced frames, such as TR_1 and TR'_1 TR_2 and TR'_2, each consisting of a succession of horizontal scanned lines, each line (such as $l_{1.1}$, $l_{1.2}, \ldots, l_{1.17}$..in TR_1 and $l_{2.1}$ in TR_2) being constituted by a succession of elementary signals of pixels or image-points PI representative of points (such as $a_{1.1}$, $a_{1.2}$ and $a_{1.3}$ for line $l_{1.1}$; $a_{2.1}$, $a_{2.2}$ for line $l_{1.2}$; $al_{17.1}$ and $a_{17.2}$ for line $l_{1,17}$; $al_{1.1}$ and $a_{1.2}$ for line $l_{2.1}$) of scene 13a monitored by camera 13; this is why the drawing shows S(PI), representing a signal S composed of pixels PI.

In the known manner, S(PI) comprises frame synchronization (ST) signals at the beginning of each frame, and line synchronization (SL) signals at the beginning of each line.

Therefore, it can be seen that the S(PI) signal is composed of:

- a succession of sequences (successive frames) within the framework of a time domain and
- in each sequence (in each frame) by a series of sub-sequences (lines, pixels) within the framework of a spatial domain.

In the time domain, the expression "successive frames" will be used to refer to successive frames of the same kind (i.e. odd frames such as TR_1, or even frames such as TR'_1) in pairs of frames such as TR_1-TR'_1, forming successive images of the digital video signal S(PI), and the expression "successive pixels in the same position" will denote successive values of the pixels (PI) in the same location in successive frames of the same kind, for example $a_{1.1}$ of $l_{1.1}$ in frame TR_1 and $a_{1.1}$ of $l_{2.1}$ in the next corresponding frame TR_2.

Furthermore, device 11 includes outputs 14 outputting a number of digital signals, produced by it, which are useful to signal the existence of an area or "object" (in the more general sense described above) with relative displacement and its localization, together with its speed and direction of displacement, if it is significant in an effectively stationary environment, in particular the complex signal ZH diagrammatically grouping signals signaling the existence and localization of this area or object, the speed V and the oriented direction of displacement DI, and possibly the input digital video signal S delayed to make it synchronous with the previous signals to take account of their calculation time, this delayed signal SR being used to display the image perceived by camera 13 on a monitor or television screen 10, at the time at which is available the information concerning the area in relative displacement (if any), namely the signal ZH, (V, DI), that can be used in a processing and checking assembly 10a.

We will now refer to figure 2 to describe the structure of the first part of the device 11 shown in figure 1, this first part being represented by dashed lines inside frame 11a on this figure 2.

The main components of assembly 11a are firstly a temporal processing unit 15, with an associated memory

16, and secondly a spatial processing unit 17, with an associated delay unit 18 and sequencing unit 19, and a pixel clock 20 setting the rhythm for the temporal processing unit 15 and the sequencing unit 19.

The temporal processing unit 15 which, among other functions, smoothes the video signal:

- uses the digital video signal S, originating from the video camera 13 and consisting of a succession of pixel values PI, and HP pulses, generated by clock 20 (starting from signal S) at the pixel rate in a frame (in particular 13.5 MHz), to generate a number of signals as described below with reference to figure 3, and exchanges the values of two of these signals with memory 16, namely the smoothed values L of the digital video signal and the values C of the smoothing time constant, the values L and C being followed by the letter O for values stored in memory 16 from unit 15, or the letter I for values originating from memory 16 and reaching unit 15, and
- produces a binary output signal DP, identifying whether or not a limit has been exceeded, and a digital signal CO, indicating the updated calculated value of the time constant, namely the value CO sent to memory 16.

The computing and/or comparison blocks constituting the temporal processing unit 15 are shown in detail in figure 3; unit 15 comprises four blocks 15a, 15b, 15c, 15d.

The first block 15a of unit 15, using

- the input digital video signal S, composed of a succession of pixel signals PI, and
- a smoothed value LI of this signal S for the immediately preceding corresponding frame, calculated beforehand by unit 15 as LO and temporarily stored in memory 16 (as explained below)
- calculates the absolute value AB of the difference between the input values of PI and LI for the same

pixel position (for example $a_{1.1}$, of $l_{1.1}$ in TR_1 and of $l_{2.1}$ in TR_2)

$$AB = | PI - LI |$$

at the rate imposed by the HP clock signals of clock 20.

The second block 15b is a test block:

- this block receives the above mentioned digital signal AB from unit 15a and a digital signal of the threshold value SE, which could be constant, but which usually depends on the pixel value; it is then made to vary in the same sense as said value so that it forms a gamma correction (known means of varying SE to make a gamma correction being represented by the optional block 15e shown in dashed lines); and
- it compares these two digital signals representative of AB and SE in order to determine a binary signal DP, in other words a signal that may be equal to one of two values 1 and 0 to signal respectively that AB does or does not exceed said threshold SE :
- if AB is greater than SE, DP will be assigned the value 1 in unit 15b to show that the threshold has been exceeded;
- if AB is less than or equal to SE, DP will be assigned the value 0 in unit 15b to show that the threshold has not been exceeded.

In fact when DP = 1, the difference between PI and LI, in other words between the input digital video signal and the previous smoothed digital video signal, is too great, and this difference has to be reduced by reducing the smoothing time constant, and conversely if DP = 0 this time constant has to be increased.

The third block 15c makes the required variation to the value of the time constant as a function of the value of DP:

- if DP = 1, block 15c reduces the time constant by a unit value U: CO (new value of this constant) = CI (old value of the constant) - U;
- if DP = 0, block 15c increases the time constant by the same unit value U: CO = CI + U.

Consequently, block 15c receives the binary excess signal DP, mentioned above, from block 15b, on one input, and the CI signal, which is the value of the previous time constant stored in memory 16, on another input, and reduces or increases, by the unit value U, the input time constant CI which becomes CO sent to said memory 16 to replace CI.

Advantageously, the time constant, on which convergence of smoothing depends (as a function of the time necessary for the smoothed value to reach the input value of the digital video signal), is represented by a power of 2, namely by a value 2^p, and this integer number p is the number that will be reduced or increased by one unit, in other words by 1, in block 15c; in this case U = 1 for p on figure 3,

- if DP = 1, block 15c subtracts a unit (1) from factor p in the time constant 2^p which becomes 2^{p-1};
- if DP = 0, block 15c adds a unit (1) to the factor p in the time constant 2^p, which becomes 2^{p+1}.

The choice of a time constant of the type 2^p has two advantages, firstly it corresponds to the physiology of human sight and secondly it facilitates calculations and thus simplifies the structure of block 15c.

Block 15c must also comply with two conditions, namely CO must remain between two limiting values: CO must not become negative ($CO \geq 0$) and it must not exceed a limit N ($CO \leq N$). In the special case in which CI and CO are in the form 2^p, the upper limit N is represented by an integer number n which is a maximum value for p.

The upper limit N (or *n*) may either be constant or variable; if it is variable, an optional unit 15f (in dashed lines) makes this variation of N (or *n*), for example following an order by the user. The consequence of increasing N is to increase the displacement detection sensitivity, whereas reducing N improves detection of high speeds.

Finally, the fourth block 15d receives, on a first input, the value CO of the new time constant generated in block 15c, and on a second input, the incoming digital video signal S in the form of pixel value information PI, and, on a third input, the smoothed value of the previous entering digital video signal, namely LI, from memory 16, and it calculates:

$$LO = LI + (PI - LI) / CO$$

which is delivered on its output.

In fact, the term (PI - LI) / CO represents the modification made to the smoothed value of the digital video signal, taking account of the modified value CO of the time constant and it is proportional to the algebraic difference between the effective value of the current entering pixel PI from camera 13 and its previous smoothing value LI and inversely proportional to CO.

If $CO = 2^P$, then

$$LO = LI + (PI - LI) / 2^{PC}$$

using p_0, the value of p calculated in unit 15c and which replaces previous value pi of p in memory 16.

Consequently, the temporal processing unit 15 with four calculation blocks 15a, 15b, 15c, 15d:

- receives S(PI) from video camera 13, HP clock pulses for controlling the rate of operations, and SE and N (or *n*) threshold signals;
- starting from the LI and CI incoming signals from the associated memory 16, determines updated signals LO and CO that are sent into said memory to replace LI and CI

respectively and which represent the new calculated values of the smoothed digital video signal and the time constant, respectively; and

- delivers the CO signal mentioned above and the binary threshold exceeding signal DP that it, calculated from PI, LI and SE, to the spatial processing unit 17 through the delay unit 18.

The purpose of the smoothing operation is to normalize variations of the digital value of the entering video signal for each pixel or image point, namely the variation of each PI, by reducing the variation differences and replacing, for each pixel, the successive real variable values of PI at this image point by the smoothed LO values, which are less variable than the PI values.

Thus for each incoming PI, the temporal processing unit 15, in combination with memory 16, substitutes a smoothed value LO with reduced variations, by using a binary threshold exceeding or not exceeding signal DP and a time constant signal CO that are updated and transferred to the spatial processing unit illustrated in figure 4.

Each pixel inside the area of a frame can be identified by two coordinates (usually orthogonal), along an abscissa and an ordinate, namely x and y, by assigning two indices to the pixel, namely i (line number) for the y coordinate and j (pixel number in the line) for the x coordinate. Each pixel with indices i and j has a video value (value of the amplitude of the video signal) PI_{ij}.

If we now consider the variation with time t of values of PI_{ij} for successive corresponding successive frames at times t_0, t_1, t_2, t_3, ..., separated by a period T corresponding to the image period (usually equal to two frame periods) and which may be 0.04s if the frequency of the video image signal is 25 Hz, or 0.0333s if the

frequency of the video image signal is 33 Hz, or 50 Hz for sensors with progressive image (1 frame / 1 image), the signal of a video pixel with localization indices i and j has successive values denoted PI_{ijt0}, PI_{ijt1}, PI_{ijt2}, PI_{ijt3}, ...at these instants t_0, t_1, t_2, t_3,.

Within the scope of the invention, for P_{ijt} successive smoothing values LO_{ijt}, namely LO_{ijt0}, LO_{ijt1}, LO_{ijt2}, LO_{jt3}, are substituted in the temporal processing unit 15.

For each of the successive pixels or image points with coordinates i, j, at time t, in other words P_{ijt}, its real value PI_{ijt} is replaced by a smoothed value given by the formula

$$LO_{ijt} = LI_{ijt(t-1)} + (PI_{ijt} - LI_{ij(t-1)} / CO_{ijt}$$

The time constant is preferably of the form 2^{pijt}.

Calculations carried out in unit 15, and particularly in block 15d for each time interval T separating t_1 from t_0, t_2 from t_1, t_3 from t_2, etc. make the value LO_{ijt} converge towards the value PI_{ijt}, at a speed that depends on the time constant which is variable in space (and therefore depends on i and j) and in time (and therefore depends on t) and which may be written CO_{ijt}.

In all cases, the convergence of LO_{ijt} is slower when CO_{ijt} is larger. There is no longer any smoothing if $CO_{ijt} = 1$.

Obviously, in the above formulas it would be possible to substitute the cartesian coordinates y and x for the indices i and j giving the line order number and the pixel number in the line.

Block 15a calculates $AB = | PI - LI |$, the indices i, j and t being not explicit, the magnitude AB representing the instantaneous variability of the digital

video signal S(PI) with respect to the smoothed signal LI, for each triplet i,j,t.

One of the characteristics of a device according to the invention is to perform, in the temporal processing unit 15, a smoothing of the pixel values, to determine, for each pixel, the successive values of a smoothing time constant and a binary signal identifying whether or not a threshold has been exceeded or not by the absolute value of the difference between the value of the pixel and the smoothed value of this pixel for two successive corresponding frames, to distribute in a plane matrix the numerical values at a same instant, both of said time constant and of said binary signal for the pixels in a limited portion of the frame, this portion scanning the frame in order to localize, based on local variations of these two values at this instant, an area in relative movement and to determine the speed and (oriented) direction of the effective displacement as a function of this distribution, this being done by a spatial processing assembly described below with reference to figure 4.

The convergence of smoothing is done by calculation blocks 15b and 15c which determine a variation of the value of the new incoming time constant CO (actually CO_{ijt}) so as to accelerate the convergence. This is done by comparing AB (actually AB_{ijt}) with the threshold value SE, that may have a constant value, or is preferably variable, and if it is variable it may be a function of the pixel value in order to make a gamma correction : if the pixel value increases, the limit increases, and vice versa. The result of the comparison of AB with the threshold value SE (actually SE_{ijt} if said value varies as a function of the pixel value, this variation being done in an optional calculation block 15e) generates, in block 15b, the binary displacement

signal DP with two possible values 1 and 0 as described above.

The binary signal DP arrives in block 15c to vary the value of the time constant. Consequently, the calculation unit 15c also receives the value of the incoming time constant CI from memory 16 and updates it using the new value of the time constant CO which (among others) is sent to memory 16, where it replaces the former value CI; in fact for values of CI and CO, are used : CI_{ij} and CO_{ij} at two successive instants, for example t_0 and t_1, separated by the interval T between two successive corresponding frames (either odd, or even).

Block 15c that receives DP and CI adds or subtracts a unit value (U) to the value of the time constant CI or one unit to p when CO is in the form 2^p, depending on whether the binary signal DP defines that a threshold value is exceeded (DP = 1) or not exceeded (DP = 0).

If the threshold is exceeded, the value of this constant is too high and it is reduced, and vice versa.

Block 15c also ensures that the new value of the time constant CO, deduced from CI by adding or subtracting a unit, remains between 0 (CO not negative) and a threshold value N (CO not greater than N). If these two conditions are not satisfied, block 15c does not modify the value of CI (which was actually inside the interval from 0 to N, both limits being included) and then CO = CI.

The upper limiting value N (or n) may be constant or variable; if it is variable, it must not exceed a limiting value N max. (or n max.); the variation, if any, being imposed by block 15f under the effect of a control available to the user.

As a variant, N or *n* may be made to depend on PI (where N, *n* and PI are actually assigned the three indices *i*, *j*, *t*) in order to ensure regulation of the variation of LO (which is calculated in block 15d) as a function of the level of PI, which may be expressed "N_{ijt} or n_{ijt} is a function of PI_{ijt}", the determination of N_{ijt} or $n_{ijt} = f(PI_{ijt})$ being done in a calculation block substituted for the illustrated block 15f and to which the value of PI is delivered by video camera 13, in addition to N.

Advantageously, the time constant C (actually each C_{ijt}) may be forced to satisfy the condition of having a numerical value which is either a multiple of 2, and more particularly a power of 2 regardless of the values of i, *j* and *t*; in this particular case $C_{ijt} = 2^{p(ijt)}$, where *p* is a small integer number that is a function of *i*, *j* and *t* and can be represented by a small number of bits. This condition provides the additional advantages mentioned above:

- the smoothing convergence law is similar to the physiology laws of human sight;
- the electronic construction of blocks 15c and 15d is simplified; in particular, in block 15d, which is designed to determine the variation in the value of the time constant, for each *i*, *j* pair, by the following formula

$$LO = LI + (PI - LI)/CO$$

the calculations are simplified if CO is in the form 2^p (where *p* is a small integer number), the value of the threshold value *n* of *p* itself being a small integer number represented by a limited number of bits.

In all cases, the new smoothing value LO of the incoming digital video signal S is sent into memory 16, where it replaces LI (for each pair of indices *ij*).

On inspecting figures 2 and 3, it can be observed that the temporal processing unit 15, which comprises calculation blocks 15a, 15b, 15c and 15d, and possibly block 15e and/or block 15f, and which cooperates with memory 16, determines and delivers as outputs, as explained above, the following values, for each *i*, *j*, *t* triplet.

- firstly, the updated smoothed value LO which is transferred into memory 16 to replace the previous smoothed value LI;
- secondly, two digital signals, namely:
- a binary signal DP which indicates that a given threshold has either been exceeded (DP = 1) or not exceeded (DP = 0) by the variation in the absolute value of the incoming pixel signal originating from the video camera compared with the previous smoothed pixel signal for the same point; and
- an amplitude digital signal composed of the value of the updated time constant CO;

these digital signals DP and CO are received, through the delay unit 18, by the spatial processing unit 17 which will be described later, the signal CO also being entered into memory 16, where the value CO replaces the previous value CI for the same pixel.

Therefore, it can be seen that the capacity of memory 16 for storing firstly successive values of the smoothed pixel signal and secondly the time constant, assuming that there are R pixels in a frame, and therefore 2R pixels per complete image, must be at least $2R(e+f)$ bits, where e and f are the numbers of bits assigned to a pixel signal, and to a time constant respectively. In fact, the memory capacity does not need to be much bigger; it simply needs to exceed $2R(e+f)$ bits by the number of bits necessary for its correct operation, particularly for addressing and extracting

bits from smoothed pixel signals and time constant bits as a function of indices i and j. If each video image is composed of a single frame, it is sufficient to use $R(e+f)$ bits instead of $2R(e+f)$ bits.

The outgoing signals DP_{ij} and CO_{ij}, at instant t, from the temporal processing assembly 15 are analyzed and used in a spatial processing assembly illustrated in figure 4, the assembly of figures 3 and 4 being illustrated in figure 2.

In fact, the temporal processing assembly 15 processes frame signals, whereas the spatial processing assembly 17 in figure 4 processes sub-sequences of lines and pixels in a frame.

Figure 5 diagrammatically shows the temporal processing of the successive corresponding frame sequences TR_1, TR_2, TR_3 superimposed on the figure and the spatial processing in the first of these frames, in other words TR_1, illustrating the cartesian coordinates x and y and a pixel PI with coordinates y,x, in other words indices i,j at time t_1; the successive pixels with the same indices ij in the three frames TR_1, TR_2, TR_3 are indexed ijt_1, ijt_2 and ijt_3 respectively, and they have pixel values of $PIijt_1$, $PIijt_2$ and $PIijt_3$ respectively. A plane in figure 5 corresponds to the spatial processing of a frame, whereas the superposition of frames corresponds to the temporal processing (where time t is the variable).

The spatial processing unit 17, to which a delay unit 18 is associated (which is also shown in figure 4) cooperates with a control unit 19 that is controlled by a clock 20 that delivers a clock pulse HP at each successive pixel signal (figure 2 for the assembly).

Outgoing signals DP_{ij} and CO_{ij} from temporal processing unit 15 are distributed by unit 17 into a

matrix 21 of reduced size containing a number of rows and columns much smaller than the number of lines L and number of pixels M per line, respectively, in DP_{ij} and CO_{ij} at a given instant t. In particular, the matrix may include $2l + 1$ lines along the y axis and $2m+1$ columns along the x axis (in cartesian coordinates), where l and m are small integer numbers. Advantageously, l and m are chosen to be powers of 2, where for example l is equal to 2^a and m is equal to 2^b, a and b being integer numbers of about 2 to 5 for example. To simplify the drawing and the explanation, m will be taken to be l as an example (although it can be different) and $m = l = 2^3 = 8$; in this case, the matrix 21 will have 2 x 8 + 1 = 17 rows and 17 columns.

Figure 4 shows part of the 17 rows Y_0, Y_1, ... Y_{15} Y_{16} and part of the 17 columns X_0, X_1, ...X_{15} X_{16} forming the matrix 21 of unit 17.

The objective is to distribute into matrix 21 with l + 1 rows, in particular 17 rows, and m + 1 columns, in particular 17 columns, the incoming flows of DP_{ijt} and CO_{ijt}, in other words the binary threshold exceeding signals DP and the digital amplitude signals representing the time constant CO that arriving from the temporal processing unit 15, arranged in a wider matrix distribution for a frame consisting of L (in particular 312.5) lines, and M pixels per row, in particular of the order of 250 to 800 depending on the TV standard used.

In order to distinguish the two matrices, i.e. the L x M video signal matrix and the l x m matrix of unit 17, referenced 21, the indices i and j will be used for the two coordinates of the former matrix (which will only be seen when the digital video signal is displayed on a television screen or monitor) and the indices x and

y for the two coordinates of the latter (shown in figure 4); at a given instant, a pixel with an instantaneous value PI_{ijt} is characterized, at the input of the spatial processing unit 17, by two digital signals DP_{ijt} and CO_{ijt}. The L x M matrix of these two signals moves by scanning through the much smaller $(2l+1) \times (2m+1)$ matrix 21 as explained with reference to figure 4, matrix 21 materializing $(2l+1) \times (2m+1)$ pixels corresponding to the same frame.

In this matrix 21, each pixel is defined by a row order number between 0 and 16 (inclusive), for rows Y_0 to Y_{16} respectively, and a column order number between 0 and 16 (inclusive), for columns X_0 to X_{16} respectively, in the case in which $l = m = 8$, namely $2l + 1 = 2m + 1 = 17$. In this case, matrix 21 will constitute a representation in the plane of 17 x 17 = 289 pixels, whereas the matrix of the video signal will include several tens or hundreds of thousands of pixels, or even more.

In figure 4, elongated horizontal rectangles Y_0 to Y_{16} (only four of which have been represented, named Y_0, Y_1, Y_{15} and Y_{16}) and vertical lines X_0 to X_{16} (of which only four have been represented namely X_0, X_1, X_{15} and X_{16}) have been used to illustrate this matrix 21 (of unit 17) with 17 x 17 image points or pixels having indices defined at the intersection of an ordinate row and an abscissa column. For example, the pixel position P_{88} is at the intersection of column 8 and row 8 as illustrated in the figure at e, which is the center of matrix 21.

In order to carry out the successive spatial distribution of portions of this L x M matrix within the $(2l + 1)(2m + 1)$ matrix 21, the unit 17 is associated with a delay unit 18 which receives firstly the DP and CO signals (with assigned indices ijt) and secondly the incoming pixel signal S, in other words PI (also with

assigned indices *ijt*), and also an HP signal from a clock 20 and line sequence signals SL and column sequence signals SC (figures 2 and 4).

As shown in figure 1, the S(PI) signal includes not only pixel value signals, such as $a_{1.1}$, $a_{1.2}$, making up a time sequence (successive frames) and spatial sub-sequences (pixels per line in each frame), but also synchronization signals ST, SL, starting from which the clock unit 20 deduces a clock signal for example at a frequency of 13.5 MHz, namely one mark pulse for each pixel in a video frame, and also blanking signals BL that make unit 19 non-operational during the above mentioned synchronization signals.

In response to these HP and BL signals from clock 20 (Figure 2), the rate control unit 19 delivers a line sequence signal SL at a frequency equal to the quotient of 13.5 MHz by the number of columns per frame (for example 400) to the delay unit 18, together with a frame signal SC, the frequency of which is equal to the above mentioned quotient 13.5/400 MHz divided by the number of rows in the video image, for example 312.5, and HP.

Unit 18 uses these SL and SC signals and the HP clock signal to control the row by row spatial distribution in matrix 21.

Consequently, the successive rows Y_0 to Y_{16} receive the DP and CO signals :

- not delayed (row Y_0);
- delayed by one TP period, equal to the duration of a frame line (row Y_1);
- delayed by 2 TP (row Y_2);

and so on until

- delayed by 16 TP (row Y_{16}).

The successive delays having the duration of a frame row, namely TP, are carried out in a cascade of sixteen delay circuits r_1, $r_2 \ldots r_{16}$ that serve rows Y_1,

$Y_2 \ldots Y_{16}$ respectively, row Y_0 being served directly by the DP and CO signals without any delay arriving from unit 15.

All circuits r_1, r_2, $\ldots r_{16}$ may be built up by a delay line with sixteen outputs, the delay imposed by any section thereof between two successive outputs being constant and equal to TP.

The rate control unit 19 controls the scanning of the entire L x M frame matrix over the $(2l + 1)(2m + 1)$ matrix 21, in rotation for successive rows of successive frames making use of the row sequence signals SL, using the following procedure.

The circular displacement of pixels in a row of the frame matrix on the 17 x 17 matrix, for example from X_0 to X_{16} on row Y_0, is done by a cascade of sixteen shift registers d on each of the 17 rows from Y_0 to Y_{16} (giving a total of 16 x 17 = 272 shift registers) placed in each row between two successive pixel positions, namely the register d_{01} between positions PI_{00} and PI_{01}, register d_{02} between positions PI_{01} and PI_{02}, etc. Each register imposes a delay TS equal to the time difference between two successive pixels in a row or line, using the column sequence signals SC.

Note that, due to the fact that rows l_1, $l_2 \ldots l_{17}$ in a frame TR_1 (figure 1), for S(PI) and for DP and CO, reach unit 18 shifted by TP (complete duration of a row) one after the other and that unit 18 distributes them with gradually increasing delays of TP onto rows Y_0, $Y_1 \ldots Y_{17}$ these rows display the DP and CO signals at a given time for rows l_1, l_2, $\ldots l_{17}$ in the same frame portion.

Similarly in a given row such as l_1, successive pixel signals $a_{1.1}$, $a_{1.2}$...arrive shifted by TS and the shift registers $\underline{d}$ impose a delay also equal to TS;

therefore, the result is that the DP and CO signals for pixels in a given row Y_0 to Y_{16} in matrix 21, pixels available on this row, are contemporary, in other words they correspond to the same frame portion.

Therefore, both for the lines and for the pixels in these lines within a portion of the frame, the objective is to carry out a purely spatial processing since matrix 21 contains in its 17 x 17 pixel positions, the values of DP and CO for the 17 pixels in each of the 17 rows in the matrix of the same digital video signal S (PI), although these pixels such as $a_{1.1}$ arrive successively line by line and pixel by pixel in each of the successive rows (figure 1) in unit 18, with the corresponding DP and CO signals that are displayed.

The signals representing the contemporary COs and DPs of matrix 21 are available at a given instant on the 16 x 17 = 272 outputs of the shift registers, as well as upstram of the 17 registers ahead of the 17 rows, i.e. registers $d_{0.1}$, $d_{1.1}$....$d_{16.1}$, which makes a total of 16 x 17 + 17 = 17 x 17 outputs for the 17 x 17 positions $P_{0.0}$, $P_{0.1}$,...$P_{8.8}$...$P_{16.16}$.

Inside matrix 21, around the center $\underline{e}$ thereof having coordinates $x = 8$, $y = 8$ (which is why the number of rows and the number of columns in matrix 21 is preferably odd, $2l+1$ and $2m+1$ respectively), we may consider in particular a small matrix containing 3 rows and 3 columns where the central element of the 9 elements thereof is effectively the pixel $\underline{e}$ with coordinates $x = 8$, $y = 8$. Let this small matrix M3 be:

$$\begin{matrix} a & b & c \\ d & \underline{e} & f \\ g & h & i \end{matrix} \quad \text{(M3)}$$

in which the central element $\underline{e}$ has been underlined.

To this 3 x 3 element matrix, containing 8 positions *a*, *b*, *c*, *d*, *f*, *g*, *h*, *i* around the central element or position *e*, correspond eight oriented directions, each starting from the central position *e* up to one of the other eight positions.

For this purpose, the eight directions may be identified using the Freeman code illustrated in figure 6, the directions being coded 0 to 7 starting from the x axis, in steps of 45°. In the Freeman code, the eight possible oriented directions, numbered 0 to 7, can be represented by a 3-bit number since $2^3 = 8$, giving eight possibilities.

Considering the small matrix M3 mentioned above, the 8 directions, according to the Freeman code, starting from the central position *e* are as follows:

3	2	1
4	*e*	0
5	6	7

as shown on figure 6.

Returning to matrix 21 in figure 4 with 17 x 17 image points or pixels, we will now describe how to identify an area in relative movement with respect to an approximately motionless environment in the scene observed by video camera 13 and therefore represented in the video digital signal S composed of pixels PI_{ijt}, and how to determine the speed and oriented direction of an effective displacement with respect to an approximately motionless environment.

Between two successive frames, such as TR_1 and TR_2 (figure 5), the PI_{ij} pixels in the signal S will be characterized, concerning their variation between instant t_1 (first frame) and instant t_2 (second frame), by the two signals DP_{ij} and CO_{ij} (distributed by scanning in matrix 21.

There is a significant variation of the pixel value at a point in this matrix if DP = 1 for this point. Therefore, an area effectively moving is identified by the matrix area within which DP = 1 at each point.

In fact a calculation unit 17a is used for examining at the same time the various nested square matrices centered on *e*, with dimensions 15 x 15, 13 x 13, 11 x 11, 9 x 9, 7 x 7, 5 x 5 and 3 x 3, within the 17 x 17 position matrix 21, the 3 x 3 matrix being the M3 matrix mentioned above. The device determines which of these matrices is the smallest in which DPs = 1 and are aligned along a straight line which determines the direction of movement of the area in which DPs = 1 and which determines a variation of +1 and -1 around CO. For a variation of +1 and -1 around CO, DP must be equal to 1 at each value if the test is to be satisfied. The smallest matrix participating in the test is chosen (principal line).

Then, within this moving area in one of the nested matrices, for example in the small matrix M3 with 3 x 3 elements, the device determines if CO varies on each side of the central position in a given direction, from +1 in an oriented direction and -1 along in the opposite oriented direction. For example, if we have -1, 0, +1 in the (oriented) direction 1, in other words in positions *g*, *e*, *c* respectively in the small matrix M3, then the displacement exists in this matrix from right to left in the (oriented) direction 1 in the Freeman code (figure 6). Obviously, at the same time in this direction in the small matrix : DP = 1. The displacement speed is higher when the matrix, among the 3 x 3 to 15 x 15 nested matrices, in which CO varies from +1 or -1 between two adjacent positions along a direction is larger. For example, if we have -1, 0, +1 in oriented direction 1, in other words *g*, *e*, *c* in the 9 x 9 matrix denoted M_9, the

displacement will be faster than if we have -1, 0, +1 in the 3 x 3 matrix denoted M_3 (figure 7).

Since CO is a power of 2 and is represented by this power in the preferred embodiments, an extended range of speeds can be identified by using only a few bits for the power of 2, while identifying a relatively low speed (that can be chosen by increasing the variation or margin for the test, for example -2, 0, +2 in positions *g*, *e*, *c* in the 3 x 3 matrix M3 would indicate a speed half as fast as the speed corresponding to -1, 0, +1 for the same positions in matrix M3).

Two tests should also be foreseen to remove uncertainties:

- a first test chooses the largest variation, in other words the highest time constant, if there are variations of CO along several directions in one of the nested matrices, for example in the small 3 x 3 element matrix M3;
- a second test arbitrarily chooses one of two (or more) directions along which the variation of CO is identical, for example by choosing the smallest value of the Freeman code; this case usually arises when the real (oriented) direction of displacement is approximately between two successive coded directions in the Freeman code, for example between directions 1 and 2 corresponding to an (oriented) direction that can be denoted 1.5 (figure 6) forming about 67.5° with the x axis direction (direction 0 in the Freeman code).

Determination of the oriented direction and speed of an area in effective movement as described above is done by calculation unit 17a (figure 4) associated with unit 17 and which receives the above-mentioned 17 x 17 outputs of matrix 21, both for CO and for DP (two outputs from 21, namely $S_{0.1}$ and $S_{0.16}$ are shown). Unit 17a processes the values of CO and DP for the successive nested matrices and uses these values to determine the

oriented direction (in the Freeman code) and the displacement speed (depending on which of the nested matrices is to be selected), possibly by applying the tests mentioned above.

The scanning of an entire frame of the digital video signal through matrix 21 takes place :

- firstly for the group of the first 17 rows or lines, rows 1 to 17, in the frame : from the left of the frame to the right considering relative movement, as shown for frame TR_2 in figure 5, i.e. from portion TM_1 at the extreme left, then TM_2 offset by one column with respect to TM_1, until TM_M (where M is the number of pixels per frame line or row) at the extreme right;
- then similarly, from the left to the right, column by column, for lines or rows 2 to 18 in the frame; and
- so on, each time going down by one line; line 3 to 19... until the last group at the bottom of the frame, namely lines L - 16 ...L (where L is the number of lines per frame).

Considering figures 2 and 4, it can be seen that the signals delivered by units 17, 18 and 19, in other words the spatial processing assembly, are as follows:

- a signal V representing the displacement speed, based on the amplitude of the maximum variation of CO in the identified area, the value of which may be for example represented by a scale of eight integers from 0 to 7 if the speed is in the form of powers of 2, and therefore contains 3 bits;
- a signal DI representing the direction of this displacement, deduced from the direction of this maximum variation, the value of DI being also possibly represented by one of the eight values 0 to 7 in the Freeman code, and therefore contains 3 bits;
- a validation signal VL specifying that the result for the speed and (oriented) direction is valid, in order to be able to distinguish a valid output with V = 0 and

DI = 0, from the lack of an output due to an incident, this signal being either 1 (valid output) or 0 (no output); therefore, only one bit is necessary for VL;

- advantageously, a time constant signal CO, for example containing 3 bits;
- (the 3 or 4 signals V, DI, VL, and possibly also CO, delivered by the spatial processing unit 17 and its associated electronics);
- a delayed video signal SR composed of the input video signal S delayed in the delay unit 18 by 16 consecutive line durations TR and therefore by the duration of the distribution of the signal S in the 17 x 17 matrix 21, in order to obtain a digital video signal contemporary of the matrix representation in matrix 21, the contents of this signal possibly being displayed in clear on a television set or monitor screen;
- all three signals delivered by unit 19, namely clock HP, line sequence SL and column sequence SC signals.

Nested hexagonal matrices (figure 8) or an inverted L-shaped matrix (figure 9) may be substituted for the nested rectangular matrices in figures 4 and 7.

In the case shown in figure 8, the nested matrices (in which only the most central matrices MR1 and MR2 have been shown) are all centered on point MR0 which corresponds to the central point (in which the binary signal is "0") of matrices M3, M9 in figure 7. The advantage of a hexagonal matrix system is that it allows firstly to use oblique coordinate axes x_a, y_a and secondly a breakdown into triangles with identical sides, to carry out an isotropic speed calculation.

The matrix in figure 9 is composed of a single row (L_u) and a single column (C_u) starting from the central position MR_u in which the two signals DP and CO respectively are equal to "1" for DP and increase or decrease by one unit for CO, if a movement occurs.

It is thus determined if the direction of (relative) displacement is:

- in the direction of the x coordinate: the CO signal is identical in all positions (boxes) in column CU, and the binary signal DP is equal to 1 in all positions in row L_u, from the origin MR_u, with the value CO_u, up to the position in which CO is equal to CO_u +1 or -1 inclusive;
- in the direction of the y coordinate: the CO signal is identical in all positions (boxes) in row L_u, and the binary signal DP is equal to 1 in all positions in column C_u, from the origin MR_u, with the value CO_u, up to the position in which CO is equal to CO_u +1 or -1 inclusive;
- or finally oblique relative to x and y : the binary signal DP is equal to 1 and CO is equal to CO_u in positions (boxes) of L_u and in positions (boxes) of C_u, the slope being determined by the perpendicular to the line passing through the two positions in which the signal CO_u changes by the value of one unit, the DP signal always being equal to 1.

Figure 9 shows the case in which DP = 1 and CO_u changes value by one unit in the two specific positions L_{u3} and C_{u5} and indicates the corresponding slope P_p.

In all cases, the displacement speed is a function of the position in which CO changes value by one unit.

If CO changes by one unit in L_u or C_u only, it corresponds to the value of the CO variation position.

If CO changes by one unit in a position in L_u and in a position in C_u, the speed is proportional to the distance between MR_u and E_x (intersection of the line perpendicular to C_u-L_u passing through MR_u).

The system described up to now with reference to figures 1 to 9 is advantageously complemented by a supplementary system which will be described with

reference to figures 11 to 16 to constitute a global system 22 illustrated in figure 10 which represents firstly the assembly 11a in figure 2 indicating the V, DI, VL, C and SR and composite F (HP, SL, SC) signals transferred by assembly 11 (which receives the incoming digital video signal S) to assembly 22a (which delivers a composite outgoing signal ZH).

The connection line $Z-Z_1$ between assemblies 11a and 22a is shown in figures 2, 10 and 11, the outputs from assembly 11a being connected to the inputs of assembly 22a along $Z-Z_1$ to transfer the above mentioned signals.

The output delivered by unit 22a and therefore from global device 22 is composed of a composite signal ZH supplying the required information on the area in relative movement in scene 13a monitored by video camera 13.

The complementary unit 22a connected to the outputs of unit 11a is illustrated in figure 11 in the form of functional blocks, remembering that this figure is connected along line $Z- Z_1$ (in the upper part) to figure 2 along line $Z- Z_1$ (in the lower part of figure 2).

The unit in figure 11 is composed essentially of a histogram formation and utilization device, and comprises:

- a micro-bus line 23 that transports a number of digital signals described in detail below;
- six histogram formation and processing blocks or units referenced 24, 25, 26, 27, 28, 29, for the histograms of the signals : delayed digital video SR, speeds V, oriented directions (in Freeman code) D1, time constants CO, first axes $x(m)$ and second axes $y(m)$, respectively;
- six linear combination blocks or units 30, 31, 32, 33, 34, 35 combining their inputs from bus 23 each to form

a validation signal V_1, V_2, V_3, V_4, V_5, V_6 for the six blocks 24, 25, 26, 27, 28, 29 respectively, 30 being associated with 24, 31 being associated with 25, 32 being associated with 26, 33 being associated with 27, 34 being associated with 28 and 35 being associated with 29;

- a moving area block or unit 36 coordinating outputs from blocks 28 and 29 for the $x(m)$ and $y(m)$ axes; and
- a datum line change block or unit 37 receiving orientation signals $x(m)_0$ and $y(m)_0$ for the $x(m)$ and $y(m)$ axes, as well as pixel clock signals HP, line sequence signals SL and column sequence signals SC (these three signals being grouped together in F bundle in figures 2, 4, 10 and 11) originating from unit 19 in figures 2 and 4, and forming the $x(m)_1$ and $y(m)_1$ signals sent to units 28 and 29 respectively.

Units 24, 25, 26 and 27 each send their SSR, SV, SDI, SDO outgoing signals to bus 23, whereas units 28 and 29 each deliver a signal $x(m)_2$, $y(m)_2$ to the two inputs of the moving area unit 36, which combines these two signals from units 28 and 29 and delivers a composite signal $xy(m)$ onto bus 23.

Operation of each of the histogram formation and processing units 24 to 29, which receive firstly a validation signal V_1, V_2, V_3, V_4, V_5 or V_6, from the associated linear combination unit 30 to 35, and secondly an SR, V, DI, CO, $x(m)_1$ or $y(m)_1$ signal to be processed as input, is the same, and this is why only the operation of one of these units, namely unit 25 for the formation and processing of speed histograms V, taken as an example, is described ; the only difference for the other analog units 24, 26, 27, 28 and 29 is the variable that is processed, noting that the different incoming signals for the six units 24 to 29 are all digital signals, which

enables a similar structure and operation of these six units.

Figure 12 diagrammatically represents the envelopes of histograms 38 and 39, respectively in x and y (cartesian coordinate axes of the 17 x 17 element matrix 21 in figure 4), of the speeds V of the area in displacement (in figure 14, elements such as C_1, C_2 of the histogram envelope 38 are shown); x_M and y_M represent the x and y coordinates of the maxima of the two envelopes 38 and 39 respectively, whereas l_a and l_c for the x axis and l_c and l_d for the y axis represent the limits of the range of significant or interesting speeds, l_a and l_b being the lower limits and l_b and l_d being the upper limits of the significant portions of the histograms.

The vertical lines L_a and L_b of abscissas l_a and l_b and the horizontal lines L_c and L_d of ordinates l_c and l_d form a rectangle that surrounds the cross hatched area 40 of significant speeds (for all x and y directions), a few microareas 41 with lower speeds, and which will be ignored, exist close to the main area 40.

Therefore, all that is necessary to characterize the area with the largest variation of the parameter shown in the histogram, the speed V in this particular case, is to identify the coordinates of the four limits l_a, l_b, l_c and l_d and the two maxima x_M and y_M. Block 25 continuously delivers this information concerning V onto the micro-bus 23.

Similar blocks 24, 26 and 27 continuously deliver information concerning the area of the maximum value for SR, DI and CO, respectively, onto this micro-bus 23 in the same way.

Finally, similar blocks 28 and 29 continuously deliver information concerning the area of the maximum

value of $x(m)_1$ and $y(m)_1$ respectively to unit 36, which combines this abscissa and ordinate information referenced $x(m)_2$ and $y(m)_2$ respectively into a composite signal $xy(m)$ that is sent through the output of unit 36 onto bus 23.

Eventually, bus 23 transfers information concerning the area containing the maximum values of SR, V, DI, CO and $xy(m)$, in other words $x(m)_1$ and $x(m)_2$, to determine if there is a moving area in the scene observed by the video camera 13, to localize this area and to determine the displacement speed and (oriented) direction.

In figure 11, the composite outgoing signal available on bus 23 is denoted ZH. The above mentioned components of this signal ZH, particularly for V and DI, in other words the speed and oriented direction of the moving area, may be displayed in digital or analog form, may produce a light signal and/or a buzzer, particularly if the speed V exceeds a limit, or they may be transmitted, by cable, optical fiber or radio relay for remote applications, to a control unit, such as unit 10a in figure 1, located close to or remote from device 11 according to the invention.

The portion of the units in figure 12 above bus 23 is used, following the point by point processing of an entire frame and the association of points in a frame, to deduce an external global value, to determine if there is an area in the observed scene in relative movement, and if so to localize it, and furthermore, if this area is actually moving, to determine the speed and oriented direction of this displacement. This area in relative movement is identified in the observation plane along two directions x, y which are not necessarily orthogonal (for example in the application in figures 15 and 16), and is

made by the portion of the units in figure 12 under bus 23.

We will now describe the structure and operation of a histogram formation and processing block, such as 25, and of its associated linear combination block, such as 31, in more detail, with reference to figures 12, 13 and 14.

Block 25 (figure 13) includes a portion 25a forming the histogram and a portion 25b forming a classifier for the histogram, these two portions operating under the control of software materialized as a portion of the integrated circuit 25c which extracts the limits I_a, I_b, I_c, I_d of the histogram (figure 11).

The classifier 25b, as well as classifiers of other histogram formation and processing blocks 24, 26, 27, 28, 29 (for the latter two through the $x(m)$ and $y(m)$ combination unit 36),delivers its output information to bus 23, and, via said bus, to the linear combination unit 31, which therefore receives information in parallel from all classifiers in units 24, 25, 26, 27, 28, 29 and which, as a function of this information, delivers or not the validation signal V_2 to unit 25.

Using software 25c, the classifier 25b determines the various classes (each containing the same number of speed values in the case considered) that will define the envelope such as 38 or 39 (figure 12).

Figure 14 shows the successive classes C_1, $C_2 \ldots C_{n-1}$, C_n, along the x axis, together with their envelope 38 for the speed V, that are determined in classifier 25b.

Figures 15 and 16 illustrate the part taken by histograms for $x(m)$ and $y(m)$ formed by units 28 and 29 and combined in unit 36 in order to obtain a slope.

For example, we have considered the case of observing a road using a video camera with digital output

located onboard a vehicle with its associated device according to the invention.

Figure 15 shows the two road sides B_g at the left and B_d at the right of the road R, as well as the slopes of the projection P_x relative to $x(m)$ done by unit 28 and the projection P_y relative to $y(m)$ done by unit 29, slopes numbered for example from 0 to 7 (using a convention other than the Freeman code).

In order to give the best precision concerning the right side B_d, in other words maximum sensitivity of the indications (for example speed) concerning this side, the projection P_x should have a slope as close as possible to the optimum slope P_0 which is perpendicular to B_d, and particularly slope 5 in the representation shown in figure 15. The maximum value of the speed histogram will therefore be obtained for the slope 5 determined by unit 28 (figure 11).

The same reasoning is applicable to the left side B_g concerning the slope of the projection P_y and therefore the part taken by of unit 29.

The unit 36 combining the optimum two slopes supplies optimization information for the two sides B_d and B_g.

Figure 16 illustrates an application to determine the optimum slope P_0 of the projection P_x, taken for example, for satisfactory driving of a vehicle Vh in continental Europe when driving on the right side of the road (case a), or in the United Kingdom when driving on the left side of the road (case b), and finally of an aircraft Va for correct landing in the middle of an airport runway (case c).

Consequently, in order to assist in driving a land vehicle (car, truck) on a road or in flying an aircraft (airplane, space shuttle) close to an airport runway, the device according to the invention further includes means of representing the right and left sides

Bd and Bg respectively of the road or the runway, and means of orienting at least one of the coordinate axes, with variable slope, so that it remains approximately orthogonal to the corresponding side (position P_0).

Up to now, we have concentrated on describing means of determining an area in effective displacement in an approximately motionless environment by means of detecting a region in which DP ≈ 1. In order to determine an area not moving in an environment in which the background is moving (for example a broken-down vehicle or a crash on a motorway), the opposite situation occurs in which areas wherein DP ≈ 0 must be localized with respect to the environment in which DP ≈ 1. Obviously in this case, speeds will be zero in one area and the concept of direction will be meaningless. Therefore, the calculations in 17a will be different.

The device according to the invention, if it is only to be used to determine an immobile area, can be simplified with the elimination of units or blocks processing speeds and directions, particularly blocks 25, 26, 31, 32 and the number of outputs from unit 11a and inputs to unit 17 can be reduced.

In the operation unit 10a of a device according to the invention, means of displaying histograms and/or values of the DP or CO signals on a monitor screen may be provided.

The SR signal, namely the delayed digital video signal, is usually applied to a television or monitor screen 10 in order to display this signal locally or remotely on the screen, at the same time as a relative movement is signaled in order to verify the nature of this relative movement. Therefore, it is only necessary to look at the television or monitor screen 10 when a area in relative displacement has been signaled, for example by a visual alarm and/or buzzer.

It is useful to facilitate identifying an area in relative movement on the screen by imposing arbitrary colors on the delayed digital video signal SR, each color or color grade representing a movement speed and/or direction.

The various units that have been described with reference to figures 2, 3, 4, 11 and 13 are each constituted by a known type of electronic circuit, particularly with microprocessors carrying out calculations and/or comparisons or using scanning signals; memories; delay units; shift registers; units forming linear histograms and associating this type of histograms in the plane; and microbuses.

The combination of these electronic circuits, in separate units 11a or 22a or in an assembly 22 composed of 11a and 22a, may be done using two very small integrated circuits or a single very small integrated circuit, with a size of the order of 10 mm x 10 mm for example using 0.7 µm technology, the set of the two interconnected integrated units or the single integrated circuit being connected through its input to the digital video output of a video camera or other observation device and through their outputs to one or several local or remote utilization devices. In one variant, if the simplified device in unit 11a only is used, this device, preferably in the form of a single integrated circuit, is placed between said digital output and one or several local or remote utilization devices.

We will now give examples that are in no way restrictive of additional applications of a device according to the invention, particularly the device according to figures 1 and 10, in other words according to the assembly shown in figures 2 and 11 as connected along line Z-Z^1.

A first additional application shown in figure 17 is constituted by the automatic framing of a person

moving in a room, for example during a video conference. Automatic framing eliminates the movements of the person moving around, which increases the definition of the image of this person observed by a video camera with a digital output and also, if the digital video signal is compressed, simplifies this compression.

Figure 17 shows a video camera 13 of the above-mentioned type that observes the person P who can move about. The digital video signal S from the video camera is transmitted by cable, optical fiber or radio relay to a television or monitor screen 10b, but is also received by the device 11 according to the invention, and the ZH output of this device acts on a unit 42 which, in response to signals received from device 11 concerning the position and movement of person P, controls the motors 43 of said camera 13 to direct the optical axis of the camera towards the person and particularly towards his face as a function of the location, speed and direction of his movement, possibly varying the zoom setting, the focal distance and/or the focusing of the camera if the person P moves forwards or backwards.

It is also possible to control movements of one or more spotlights on a person (actor, singer) on a stage, the device according to the invention centering the person at the center of the image by varying the orientation of the video camera and controlling the direction of the spotlight, for example by one or two rotatable mirrors for each spotlight.

Another example of application of the device according to the invention is illustrated in figure 18, which shows the camera 13 or other observation device that outputs a digital video signal S to a device 11 according to the invention. In this application, the camera 13 monitors part of a motorway (highway) in order to detect an unexpected stop of a vehicle, particularly

on the hard shoulder, or if a car is stopped following a collision.

Therefore, the objective in this case is to determine whether or not an object (the vehicle) is immobile in a moving environment (other vehicles), in other words to localize the area in which DP = 0 in the 17 x 17 element matrix 21. Normally, the camera 13 observes a vehicle stream that produces DP = 1 responses, with speed and displacement direction values. However, if a vehicle stops, it is identified by a DP = 0 response in its observation area.

A unit 44 receiving both ZH and SR detects the occurrence of a area in which DP = 0 in ZH and outputs an error signal NL which firstly trips a buzzer and/or a visual alarm in a device 45, and secondly controls a switch 46 which applies the video signal S (or rather the delayed video signal SR) on a television or monitor screen 10 which enables the supervisor, alerted by the buzzer and/or visual alarm, to observe the motorway when a vehicle stops or a collision occurs in order to take the necessary measures, for example depending on the reaction of the driver of the vehicle which stopped.

Using the invention, a supervisor in a monitoring station can easily control a large number of segments on a motorway (or a main road) for each of which there is a camera 13 and a device 11 according to the invention, the ZH and SR outputs from each device 11 being transmitted by cable, optical fiber or radio relay to a single supervision station at which there is a common unit 44; the observer only needs to look at the screen in 10 if there is an incident or accident signaled by the alarm, and it is unusual that several incidents and/or accidents occur simultaneously at different locations.

In particular, unit 44 may include, at its input, a rotary type switch (not shown) sending the ZH (and SR) signals from the various assemblies 13-11 located along

the motorway successively and cyclically onto the portion of this unit 44 producing the NL signal.

The same system allows to detect not only if a vehicle is stopped, or a collision causing vehicles to stop takes place, but also a reduction in the speed of the stream of vehicles (for very slow traffic), due to the reduction of the speed in the various areas in which D = 1 and conversely a vehicle traveling at very high speed, the speed in the area observed then exceeding the allowable speed limit.

Finally, figure 19 illustrates another application of the invention, namely man-machine interaction, by movement of a hand M or more particularly fingers DG inside an area SF broken down into rectangles by a coordinates system C_x and C_y.

A video camera 13 with a digital output, associated with a device 11 according to the invention as shown in figure 1, can be used to recognize the movement of a hand M and fingers DG and can be used to control a computer (for example like a "mouse") or to control some functions in a machine. For example, the assembly 13-11 could be utilized by a deaf-dumb person, using the standard deaf-dumb sign language based on hand movements, to input alphanumeric data and therefore a text into a computer without needing to use the normal keyboard; this operation, which can also obviously be carried out by a speaking person who has learnt the deaf-dumb language, allows entry of a text into a computer without needing to use the keyboard. This type of system is not sensitive to time variations and does not require precise signaling of the start and end of the gesture.

Figures 20 and 21 diagrammatically illustrate application of the invention to surveillance of a car driver to signal if he falls asleep.

In this case, video camera 13 is placed adjacent to the bodywork inside the vehicle, for example above the rearview mirror, where it can observe the driver.

The first operation is to frame the driver, in the same way as for the application in figure 17. Figure 20 diagrammatically shows the image 1C of the driver on the video screen. Firstly, the useless right and left portions are eliminated (as represented by the horizontal cross-hatching of the image), thus restricting processing to the central part of the image between these two portions.

Then, within this central portion, it suffices to monitor the unhatched area AA of figure 21 framing the head.

The relevant displacements that are detected by the device according to the invention consist of the driver blinking (signaled by vertical movements in area AA), at a rate that changes just before he falls to sleep. If his blinking rate or speed drops below a certain threshold, a buzzer sounds and the driver is woken up.

Figure 22 shows a means of getting around the limited number of p bits representing the time constant CO when it is necessary to take into account a wider range of displacement speeds.

Mallat's diagram will be used for this purpose (see article by S. Mallat "A Theory for multiresolution signal decomposition" in IEEE Transactions on Pattern Analysis and Machine Intelligence, July 1989 p. 674-693) which consists of successively breaking down the entire video image into successive halves, identified 1, 2, 3, 4, 5, 6, 7. This creates a compression that only processes portions of images. Thus with $p = 4$, in other words $2^p = 16$, we can determine a speed within a wider range.

If initially, within the framework of the entire image, the device according to the invention indicates that the speed of the mobile object (in the broad sense of the term) exceeds the maximum speed determinable with $2^p = 16$ for the time constant, all that is necessary is to use partial observed images 1, 2, 3, 4,.... until the speed of the mobile object does not exceed said maximum speed within the frame of the partial image after compression.

All that is necessary for utilizing the Mallat composition by wavelets is to insert a unit 13A (illustrated in figure 22) in the diagram in figure 1, to perform this video signal compression. For example, this unit could be composed of the "DV 601 Low Cost Multiformat Video Codec" component made by the ANALOG DEVICES, an American corporation, described in the "ADV 601 Preliminary Data Sheet", January 1996. Figure 2 shows an optional compression unit of this type 13a.

Finally, the number of bits (1, 3 (if $p = 3$), 8, 15) transported on some microbuses have been shown in figures 2 and 3, which demonstrates the possible size savings for the various functional blocks as they process a reduced number of bits.

It is quite clear that the invention is capable of detecting a relative movement in a scene observed by an opto-electronic device such as a video camera which transmits the observed scene in the form of a digital video signal composed of a succession of frames, themselves composed of a succession of lines composed of a succession of pixels, this digital signal being analyzed in order to identify an area in relative displacement, with indication of the speed and (oriented) direction of this displacement if the area is in effective movement with respect to an approximately motionless environment.

Given that the device according to the invention determines the oriented direction and the displacement speed of an object (in the broadest sense of the term), means may be combined with it to use these two parameters to deduce a future position of the objects at a given instant, and to direct in advance the input video camera 13 in direction of this future position.

Note that the results obtained with the device according to the invention in no way require that the camera should be fixed, so that the camera and the associated device may be put onboard a land, air or sea vehicle (for example for embodiment of the process illustrated in figure 16).

After a very short initialization period equal to N (of the order of 10 corresponding successive frames), the device according to the invention determines the relative displacement parameters instantaneously after the end of each frame on which the temporal and spatial processing was performed, due to the recursiveness of calculations according to the invention.

We have described a preferred embodiment of the device according to the invention and a few of its applications. Obviously, this embodiment and these applications have been given as non-restrictive examples, and many variants and adaptations that will be obvious to the expert in the field may be envisaged without going outside the scope of the invention as defined in the following claims.

For example, it would be possible to implement methods of using signals delivered in output by unit 11 in figure 2 other than those illustrated in figure 11 without going outside the scope of the invention.

As concerns the applications of a device according to the invention, they are in no way limited to those given as examples in the description above. Thus, a circuit similar to that in figure 17 may be associated

directly with a camescope or video camera to stabilize it for compensating for displacements caused by accidental movements of the user.

It would also be possible to use one or preferably several devices according to the invention, associated with one or preferably several camescopes or video cameras with a digital video output placed in a room in a building, to create an "intelligent room" by means of such a system capable of detecting and localizing the presence and movement of one or several persons in the room, analyzing this movement for security or identification purposes and/or of helping in tasks to be accomplished, for example, or of watching children in another room or customers in a supermarket.

It has been assumed especially that a video signal is used with pairs of successive interlaced frames, particularly during the discussion of the memory capacity 16, with processing of the two frames in a pair in the device according to the invention. However, it would be possible to use only one frame out of two (for example the odd frame) reducing the memory capacity, but also reducing the speed at which the required information is obtained by about half. It would also be possible to use a video camera or another observation device for which the digital output only includes one frame per image.

In some applications, it is possible to associate specialized sensors with the device according to the invention, for example one or several acceleration sensors, to be able to process additional displacement parameters.

Obviously, the invention is not limited to the particular embodiments or applications described, but it includes all variants and modifications included part of the general definition of the invention.

* C L R S *

THIS SHEET HAS BEEN REPLACED

ENGLISH TRANSLATION OF THE CLAIMS AS ORIGINALLY FILED

1. Process operating in real time for identifying and localizing an area in relative movement in a scene observed by an observation system with an output constituted by a digital video signal of the type comprising a succession of corresponding frames each composed of a succession of lines, each composed of a succession of pixels, and for determining the speed and oriented direction of the displacement, said process being characterized in that it consists in carrying out the following sequence of operations on the digital video output signal:
   - a smoothing processing on said digital video output signal using a digital time constant, the numerical value of which may be modified independently for each pixel of said output signal;
   - a memorization, on the one hand, of a frame of said output signal after smoothing and, on the other hand, of the smoothing time constant associated with said frame;
   - a temporal processing for each pixel position, consisting in determining firstly the existence and secondly the amplitude of a significant variation in the amplitude of the pixel signal between the current frame and the immediately previous smoothed and memorized frame, and generating two digital signals, the first signal being a binary, or single bit, signal with two possible values, one of which represents the existence and the other represents the absence of a significant variation between two successive frames, the value of said binary signal modifying the memorized value of said time constant in order to reduce it if said signal represents a significant variation and in order to increase it if said signal does not represent such a variation, the

This Sheet Has Been Replaced

* C L R S *

THIS SHEET HAS BEEN REPLACED

reduction or increase being done in a quantified manner, whereas the second digital signal, i.e. the amplitude signal, is a multibit signal with a limited number of bits quantifying the amplitude of this variation; and

- a spatial processing consisting of the following steps for each frame of the digital video input signal :
  - o distributing only the values of a fraction of the pixels in a frame at a given observation instant (a fraction which is scanned across said matrix throughout the duration of the frame) firstly of said binary signal and secondly of said digital amplitude signal, into a matrix with a number or rows and columns that is small compared with the number of lines and the number of pixels per line in the video signal, respectively, in order to characterize the values of the pixels,
  - o determining in this double instantaneous matrix representation, a particular area in which said binary signal has the searched value, representing the presence or absence of a significant variation, and said digital amplitude signal varies, or does not vary, by a significant value for adjacent pixels in the matrix along an oriented direction starting from an origin pixel, within the same portion of the frame and therefore at the same observation instant, and
  - o generating signals representing the existence and localization of the area in relative displacement and the relative inter-frame speed and the oriented direction of this displacement if any, relative to its environment, inferred from the instantaneous matrix distribution of these two binary and amplitude digital signals.

This Sheet Has Been Replaced

* C L M S *

THIS SHEET HAS BEEN REPLACED

2. Process according to claim 1 characterized in that it also consists of:

   - o forming firstly histograms of the signal values distributed in matrices, and secondly histograms of the inclinations of two coordinate axes, with variable slope, in a plane,
   - o identifying a domain in which there is a significant variation of the processed values in each histogram formed, and
   - o for each identified domain, deducing whether or not there is an area in relative movement, and if so its localization, speed and oriented direction.

3. Process according to claim 1 or 2, characterized in that said matrix is a square matrix with the same odd number of rows and columns (2l + 1), and nested matrices containing 3 x 3, 5 x 5, 7 x 7, ...(2l + 1) x (2l + 1), elements, centered on the center of this square matrix, are considered in order to determine the smallest nested matrix in which said digital signal varies along an oriented direction starting from said center, the value of said binary signal showing that the threshold along this direction has been exceeded.

4. Process according to claim 1 or 2, characterized in that said matrix is a hexagonal matrix, and nested hexagonal matrices of increasing size, centered on the center of this hexagonal matrix, are considered, in order to determine the smallest nested matrix in which said digital signal varies along an oriented direction.

5. Process according to claim 1 or 2, characterized in that said matrix is an inverted L matrix with a single row and a single column, and nested matrices containing 3 x 3 pixels, 5 x 5 pixels, 7 x 7 pixels,.... (2l + 1) x (2l + 1) pixels are considered

This Sheet Has Been Replaced

THIS SHEET HAS BEEN REPLACED

for the single line and the single column, in order to determine the smallest matrix in which the signal varies along an oriented direction, namely the line with the steepest slope and constant quantification.

6. Process according to any one of the previous claims, characterized in that said time constant is of the form 2^p, where p is a number less than 16, and which can then be expressed in not more than 4 bits, the time constant being reduced or increased by subtracting or adding one unit to p.

7. Process according to claim 6, characterized in that successive decreasing portions of complete frames are considered using the Mallat time-scale algorithm and the largest of these portions, which gives displacement, speed and orientation indications compatible with the value of p, is selected.

8. Device operating in real time for identifying and localizing an area in relative movement in a scene observed by an observation system with an output constituted a digital video signal of the type comprising a succession of corresponding frames, successive lines in each corresponding frame and successive pixels in each line, and for determining the speed and oriented direction of this displacement, by using the process described above, said device receiving said video output signal as input and being characterized in that it comprises in combination:

- means(15) for smoothing said digital video output signal making use of a digital time constant (CO), the numerical value of which may be modified independently for each of said output signal pixels;
- means (16) for memorizing firstly a frame of said output signal after smoothing (LI) and secondly the smoothing time constant (CI) associated with said frame,

This Sheet Has Been Replaced

THIS SHEET HAS BEEN REPLACED

- a temporal processing unit (15) for analyzing the variations with time of the amplitude of the pixel signal, for the same pixel position, between the current frame and the immediately preceding smoothed and memorized frame of said digital signal, said unit comprising a memory unit (16) capable of receiving, storing and retrieving information about the corresponding previous smoothed frame, together with comparison means (15a) to determine if the absolute value of the difference, between the current pixel signal and a representative value of the pixel signal for the same pixel position in the previous frame stored in said memory unit, exceeds or does not exceed a threshold, for each position of the pixel in the input video signal frame, by generating a binary, or single bit, signal (DP) with two values, one of these two values showing that the threshold is exceeded and the other showing that it is not exceeded, and computing means (15c) capable of determining a multibit digital amplitude signal with a small number of bits, the value of which depends on the amplitude of the variation in the value of the same pixel between the current frame and the immediately preceding, smoothed and memorized, frame of the digital video signal; and
- a spatial processing unit (11), into which the outputs from said successive binary and digital amplitude signals for the pixels in a given frame produced by the temporal processing unit are input, said spatial processing unit comprising means for characterizing the amplitude values of the pixels, these means distributing only said binary and digital amplitude signals relative to the same instant, i.e. to a single frame, into a matrix (21) with fewer rows and columns than the number of lines and the number of pixels per line respectively in a

This Sheet Has Been Replaced

THIS SHEET HAS BEEN REPLACED

frame of said digital video signal, the frame being scanned across said matrix throughout the duration of a frame, spotting means for determining an area of pixels within said matrix in which, at this instant, the binary signal has the searched value, and means of determining an area of pixels in said matrix, into which, at the same instant, the digital amplitude signal for adjacent pixels varies by a significant quantity, and means which, in response to the indications from the two preceeding means, generate signals representing this area of pixels and therefore the existence and localization of an area with relative movement within the observed scene and the relative inter-frame speed and oriented direction of this area when it actually moves relative to its environment.

9. Device according to claim 7, characterized in that said spatial processing unit includes first delay means (r) in cascade, each of which imposes a delay equal to the time difference between two successive lines and second delay means (d) in cascade, for each line, each importing a delay equal to the time difference between two successive pixels in a line, the outputs from all said second delay means (d) and the input to the cascade of said second delay means (d) on each line delivering at a given instant the values of said binary signal and said digital amplitude signal at the same instant, to said identification means (17a).

10. Device according to claim 8 or 9, characterized in that it also comprises means (24-29) for forming histograms of the output values from said spatial processing unit, and histograms of the inclinations of two, variable slope, coordinates axes within a plane, means for identifying a domain, in each histogram, in which there is a significant variation in the

This Sheet Has Been Replaced

THIS SHEET HAS BEEN REPLACED

processed value, in order to validate this domain at the output, and to deduce, for all histograms, output signals that identify and localize an area of relative movement in the observed scene, if any, and the speed and oriented direction of this movement if said area is effectively moving with respect to its environment.

11. Device according to claim 8, 9 or 10, characterized in that said smoothing means (15c, 15d) include an input receiving said digital video signal S(PI) and derive, for each successive pixel in a frame of said video signal, a smoothed signal (LO) in which the temporal variations of the input digital video signal are reduced by using a threshold signal received on another input and a time constant (CO) related to each pixel position in a frame, the value of which is modified successively so that the smoothing maintains, although reduced, the trend in the variations of the incoming digital video signal, said smoothing means working in cooperation with said memory unit (16) that receives, stores and retrieves the updated values of the smoothed signal and said time constant, for each portion of pixels in a frame, and output, for each pixel position, at least the succession of values of the updated time constant and of the values of the binary signal that shows whether or not said threshold is exceeded by the absolute value of the difference between the value of the pixel and its smoothed value.

12. Device according to any one of claims 8 to 11, characterized in that said spatial processing unit (17, 18) that performs the matrix distribution, into a reduced number of rows and columns, of the outputs from said smoothing means, namely the successive values of the time constant (CO) and said binary signal, comprises identification means (17a) to identify in said matrix distribution, an area of pixels in which simultaneously either the value of

This Sheet Has Been Replaced

THIS SHEET HAS BEEN REPLACED

said binary signal corresponds to a threshold being exceeded and said time constant varies between adjacent pixels by a significant value in one direction and for producing output signals indicating the localization of said area and the speed and oriented direction of displacement in this area, or the value of said binary signal corresponds to a threshold not being exceeded and said time constant does not vary between adjacent pixels.

13. Application of the device according to claim 10, 11 or 12 to assist in driving a land vehicle on a road, or in flying an aircraft close to an airport runway, characterized in that the device also includes means for representing the right and left sides Bd and Bg of the road or respectively the runway, and means for orienting at least one of the coordinate axes at variable slope, so that it remains approximately orthogonal to the corresponding side (position P_o).

This Sheet Has Been Replaced

ABSTRACT

PROCESS AND DEVICE OPERATING IN REAL TIME FOR IDENTIFYING AND LOCALIZING AN AREA IN RELATIVE MOVEMENT IN A SCENE, AND FOR DETERMINING THE SPEED AND DIRECTION OF THE DISPLACEMENT

Application made by CARLUS MAGNUS Limited.
Inventor: Patrick PIRIM

The purpose of this invention is a process and a device operating in real time for identifying and localizing an area in relative movement in a scene and determining the speed and oriented direction of this relative movement in real time.

The process carries out temporal processing of the digital video input signal S(PI) that consists in deducing, from the variations between the value of each pixel in a frame and in the immediately preceding frame, a binary signal DP identifying whether or not there was a significant variation and a digital signal CO representing the magnitude of this variation, and a spatial processing consisting in distributing in a matrix these two signals successively for a single frame that is scanned through the matrix, and deducing the required relative movement and its parameters from this matrix distribution. The device achieves this by including a temporal processing unit 15 associated with a memory 16 and a spatial processing unit 17 associated with a delay unit 18; the clock unit 20 and the control unit 19 control the rate of operation of units 15 and 17.

(Fig. 2)

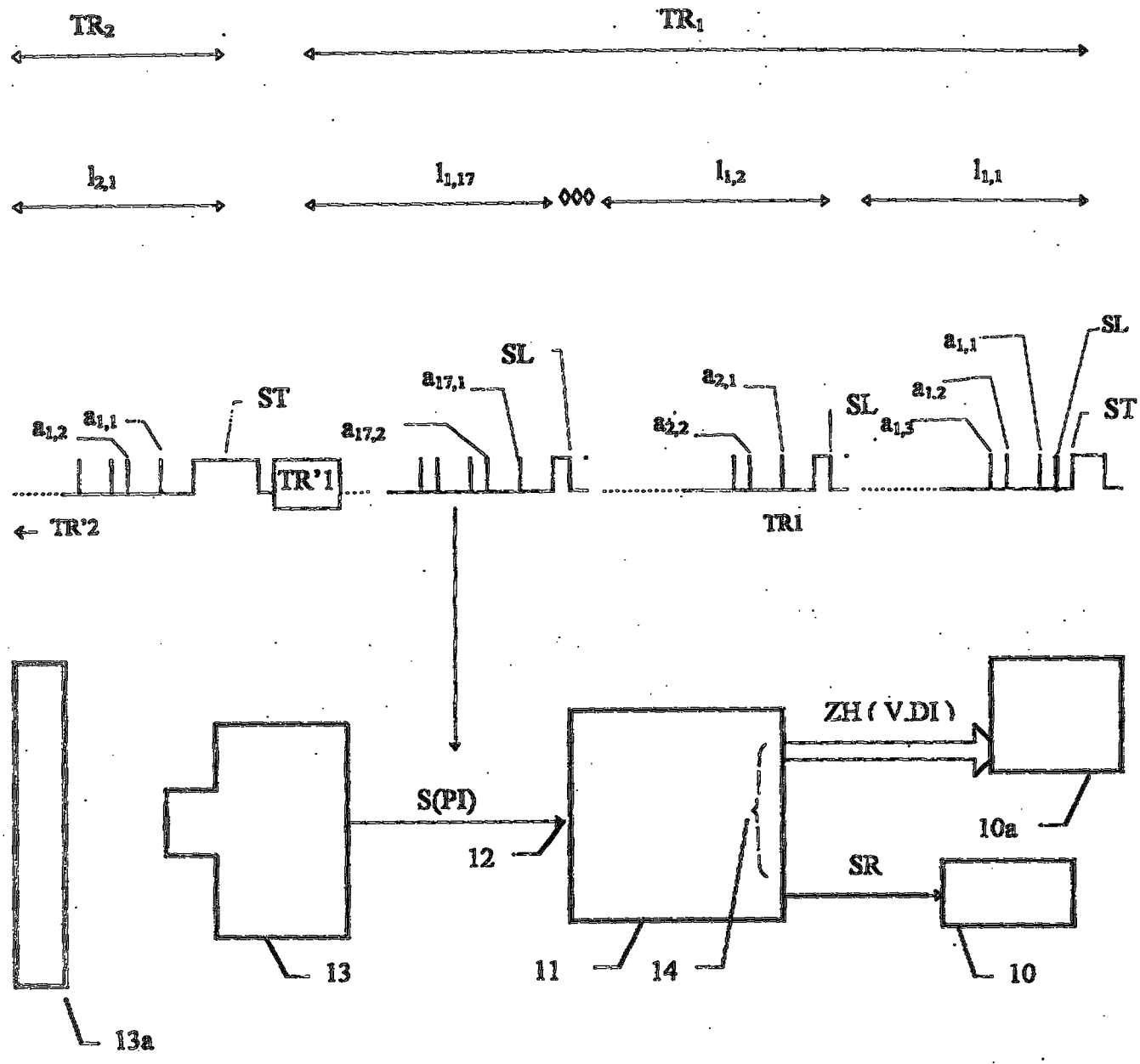

TR2
TR1
l2,1
l1,17
l1,2
l1,1
a1,2
a1,1
ST
TR'1
a17,1
SL
a17,2
a2,1
a2,2
SL
a1,1
a1,2
a1,3
SL
ST
TR'2
TR1
ZH (V.DI)
S(PI)
12
10a
SR
13
11
14
10
13a

FIG. 1

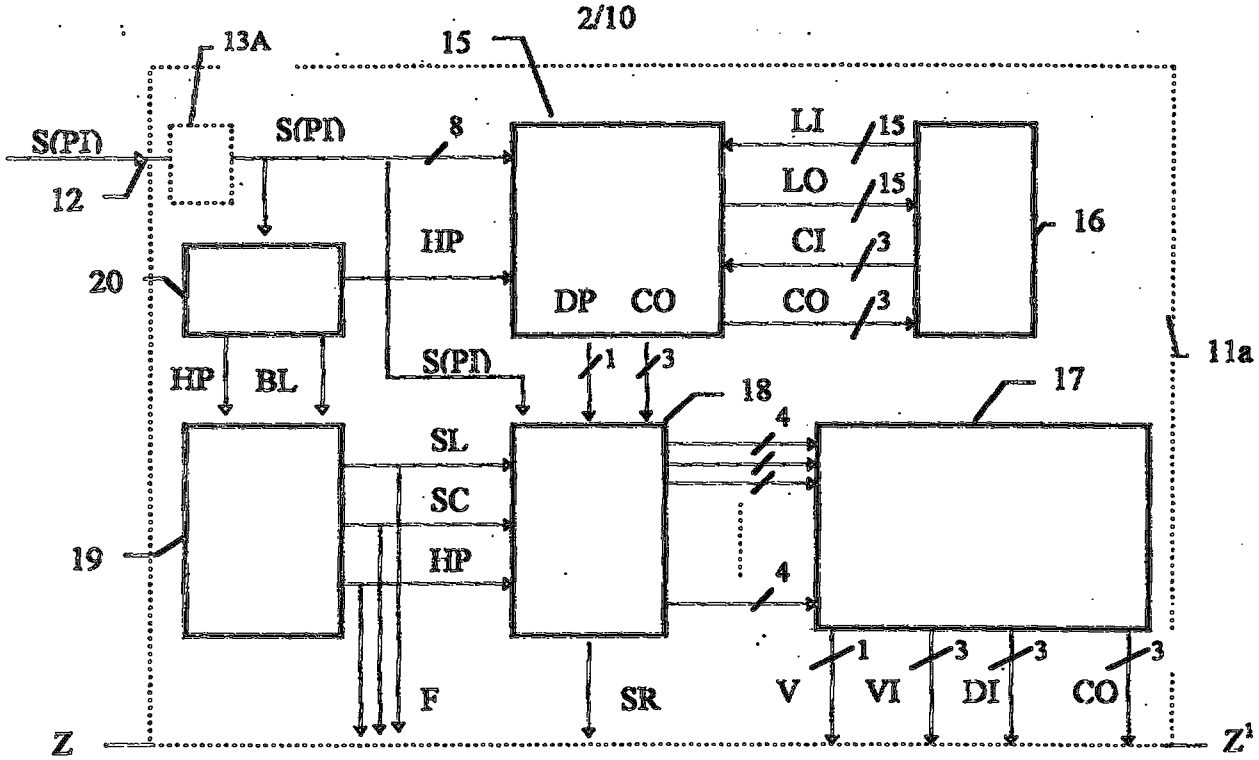
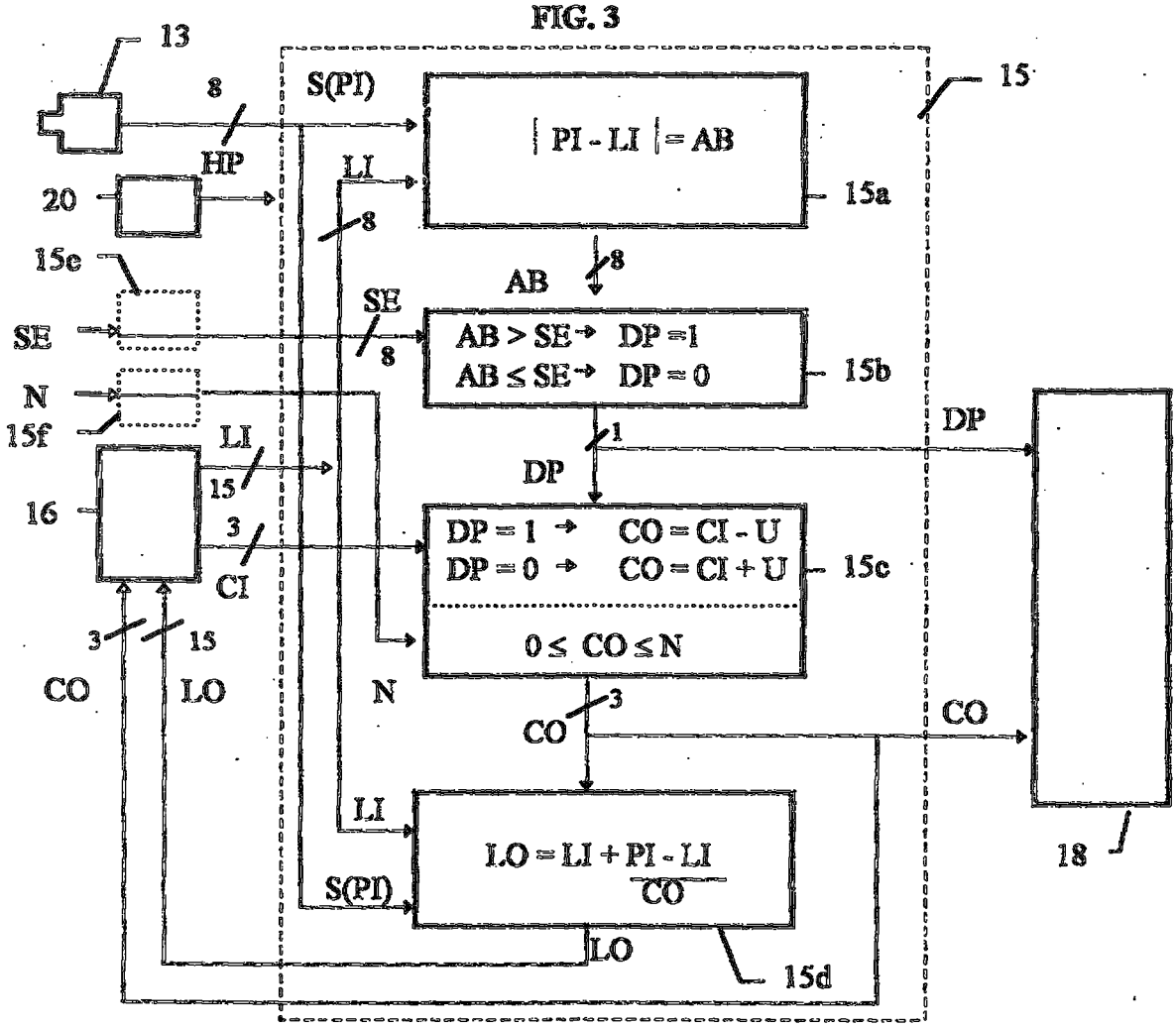

13A
15
S(PI)
12
S(PI)
8
HP
20
DP CO
LI 15
LO 15
CI 3
CO 3
16
11a
HP
BL
S(PI)
1
3
18
17
SL
SC
HP
4
4
19
F
SR
V 1
VI 3
DI 3
CO 3
Z
Z1

FIG. 2

FIG. 3


13
8
HP
20
15e
SE
N
15f
16
LI
15
CI
3
CO
3
15
LO
S(PI)
8
| PI - LI | = AB
15a
LI
8
AB
8
SE
8
AB > SE → DP = 1
AB ≤ SE → DP = 0
15b
DP
1
DP
DP = 1 → CO = CI - U
DP = 0 → CO = CI + U
15c
0 ≤ CO ≤ N
N
CO
3
CO
LI
LO = LI + PI - LI
CO
S(PI)
LO
15d
15
DP
CO
18

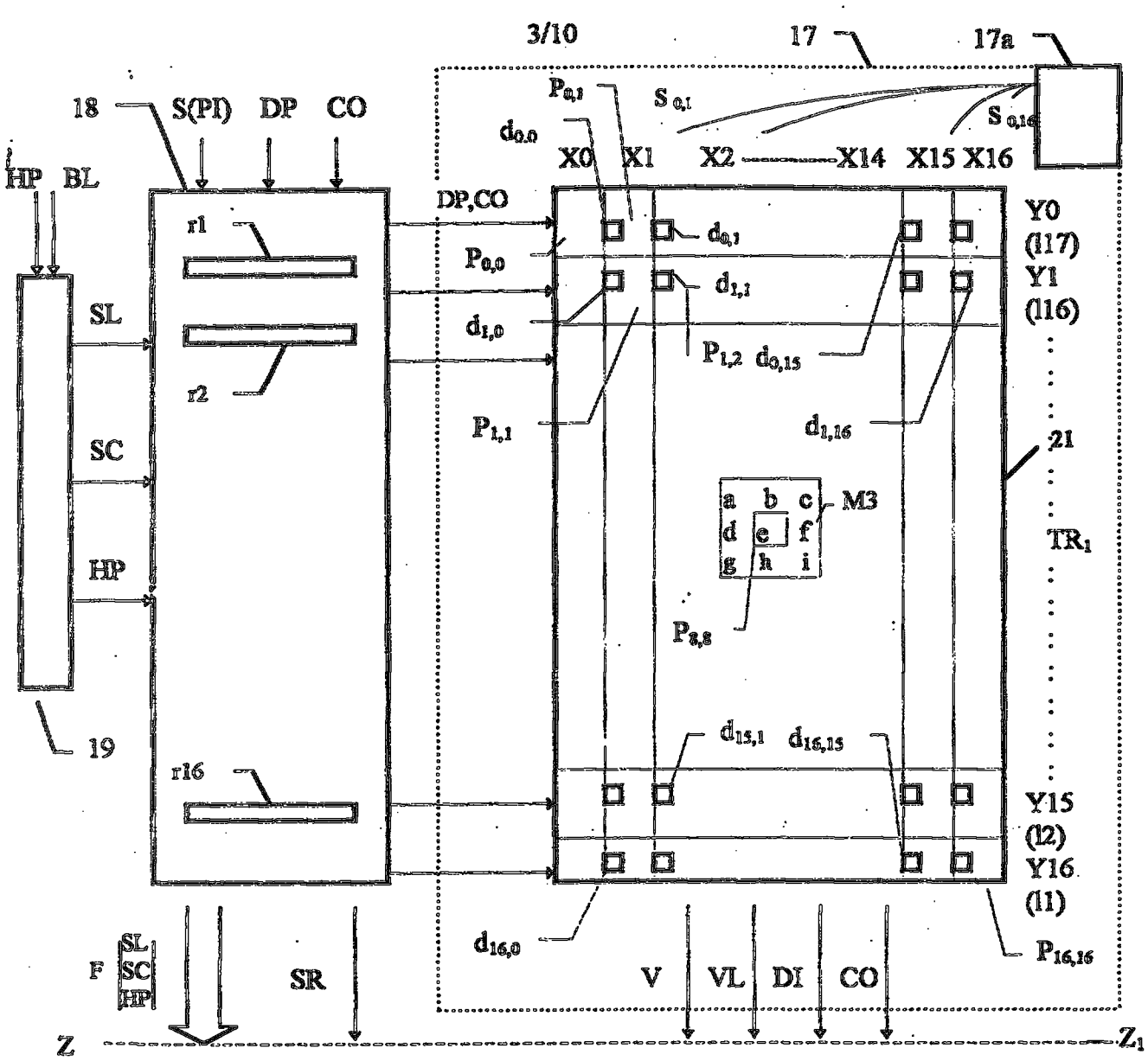


FIG. 4

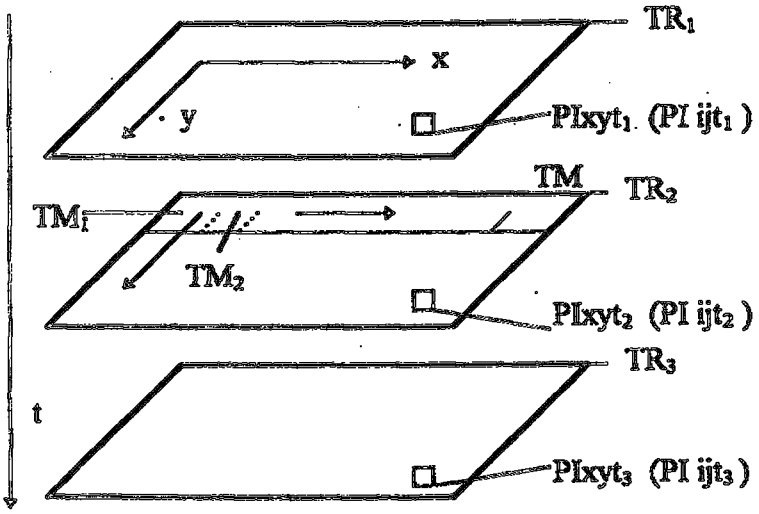


FIG. 5

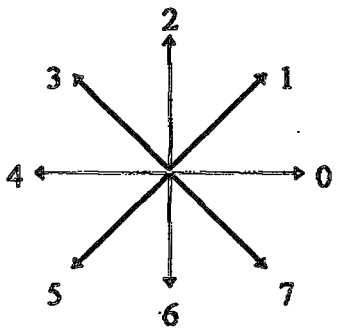
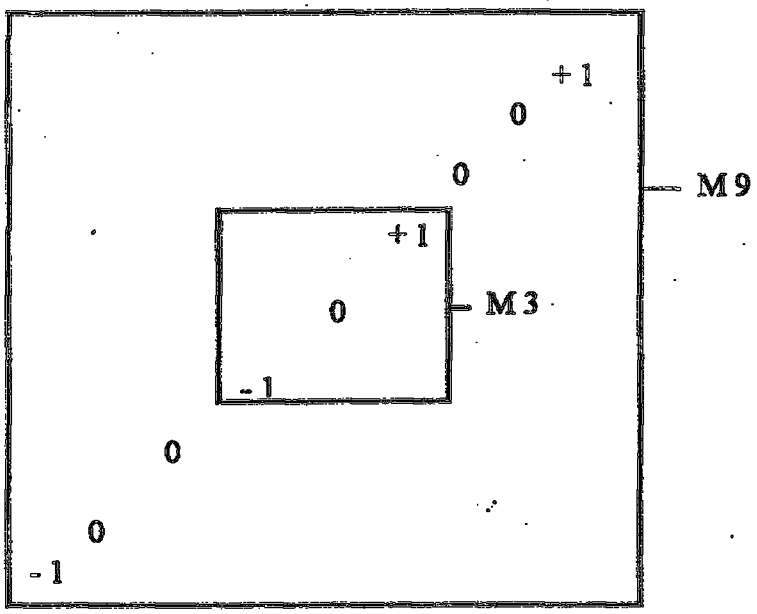


FIG. 6

FIG. 7



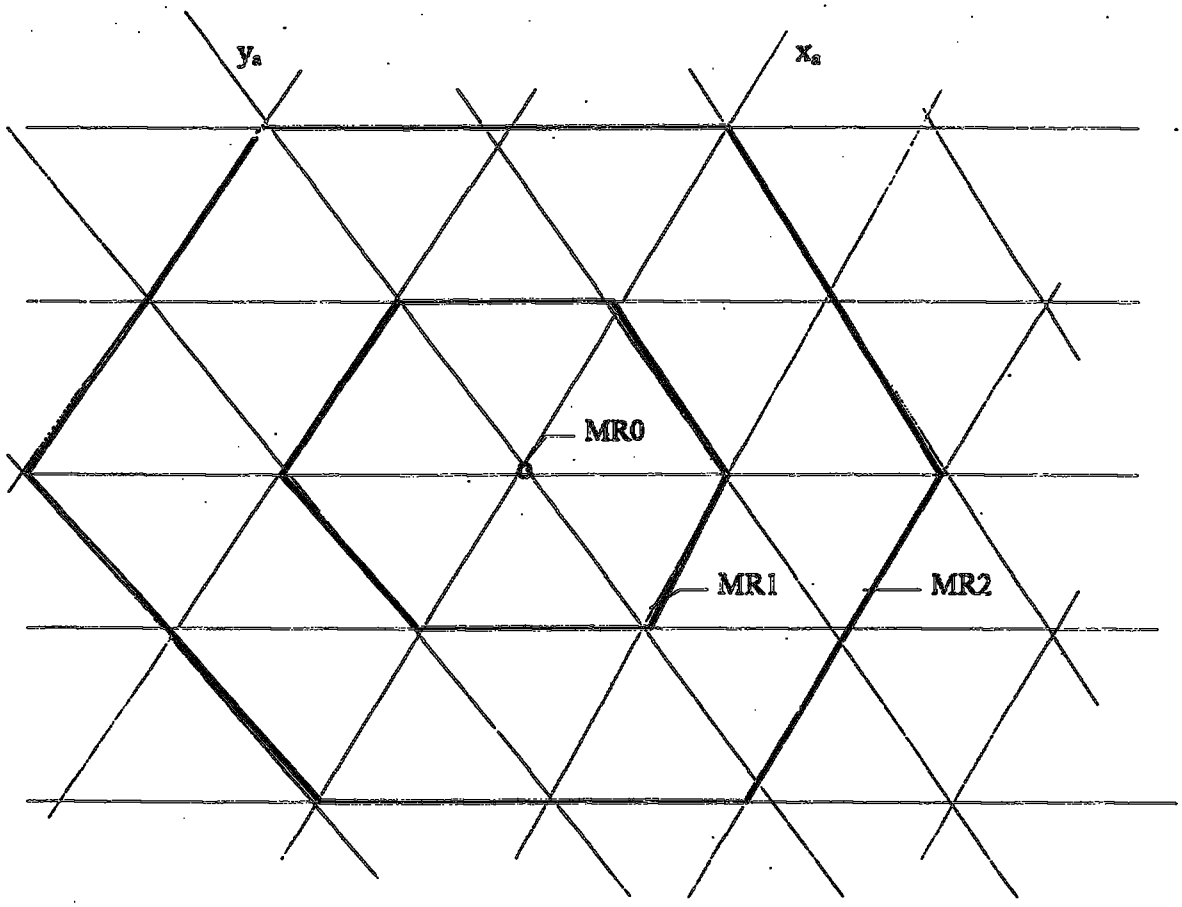


FIG. 8

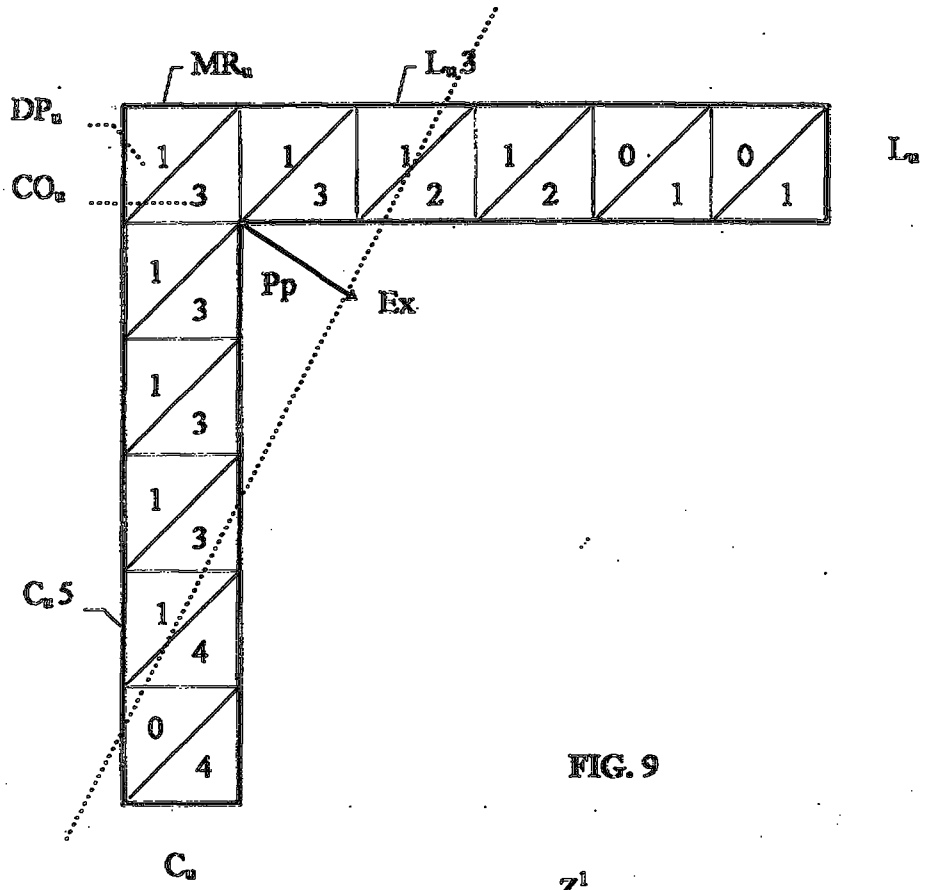

MRu
Lu 3
DPu
COu
Lu
1
3
1
3
1
2
1
2
0
1
0
1
Pp
Ex
Cu 5
1
4
0
4
Cu

FIG. 9

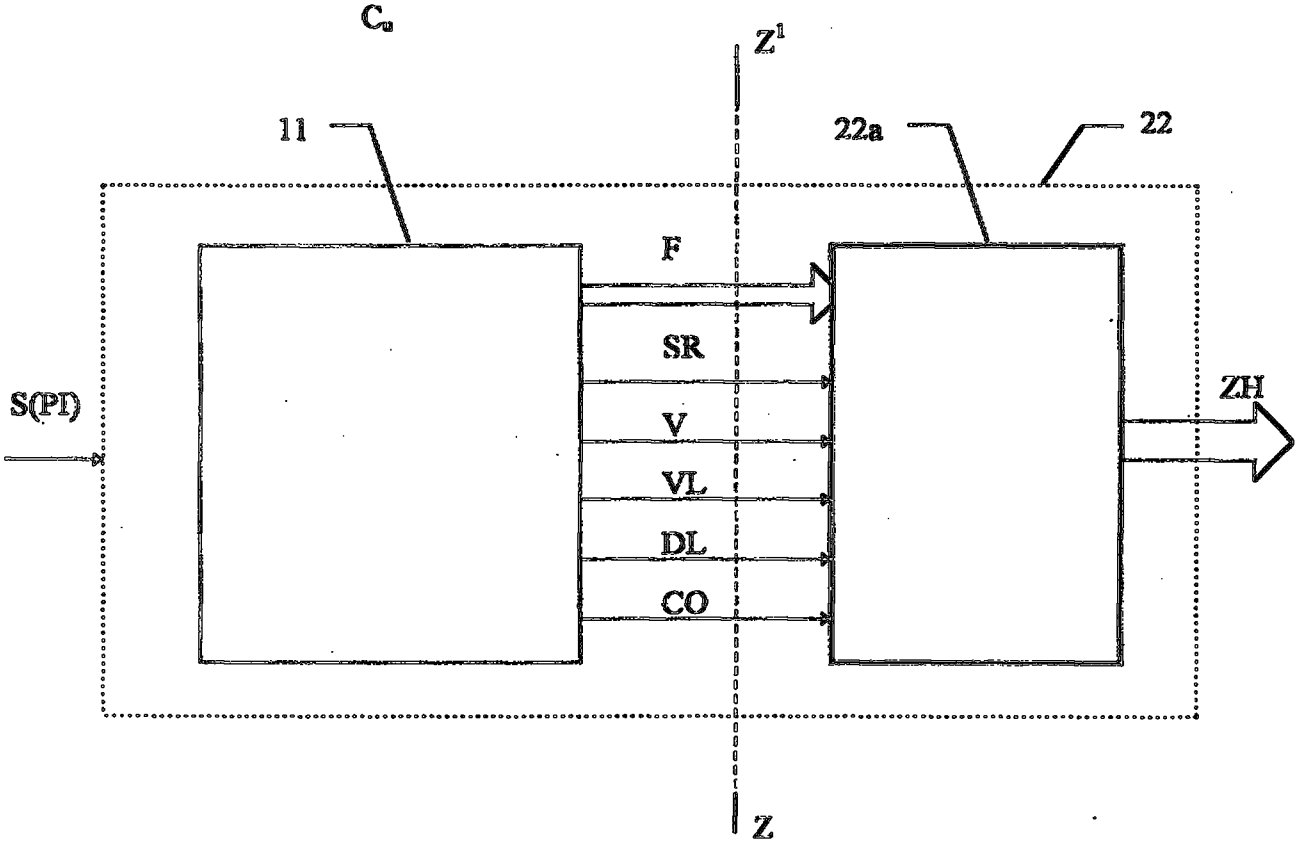

Z1
11
22a
22
F
SR
V
VL
DL
CO
S(PI)
ZH
Z

FIG. 10

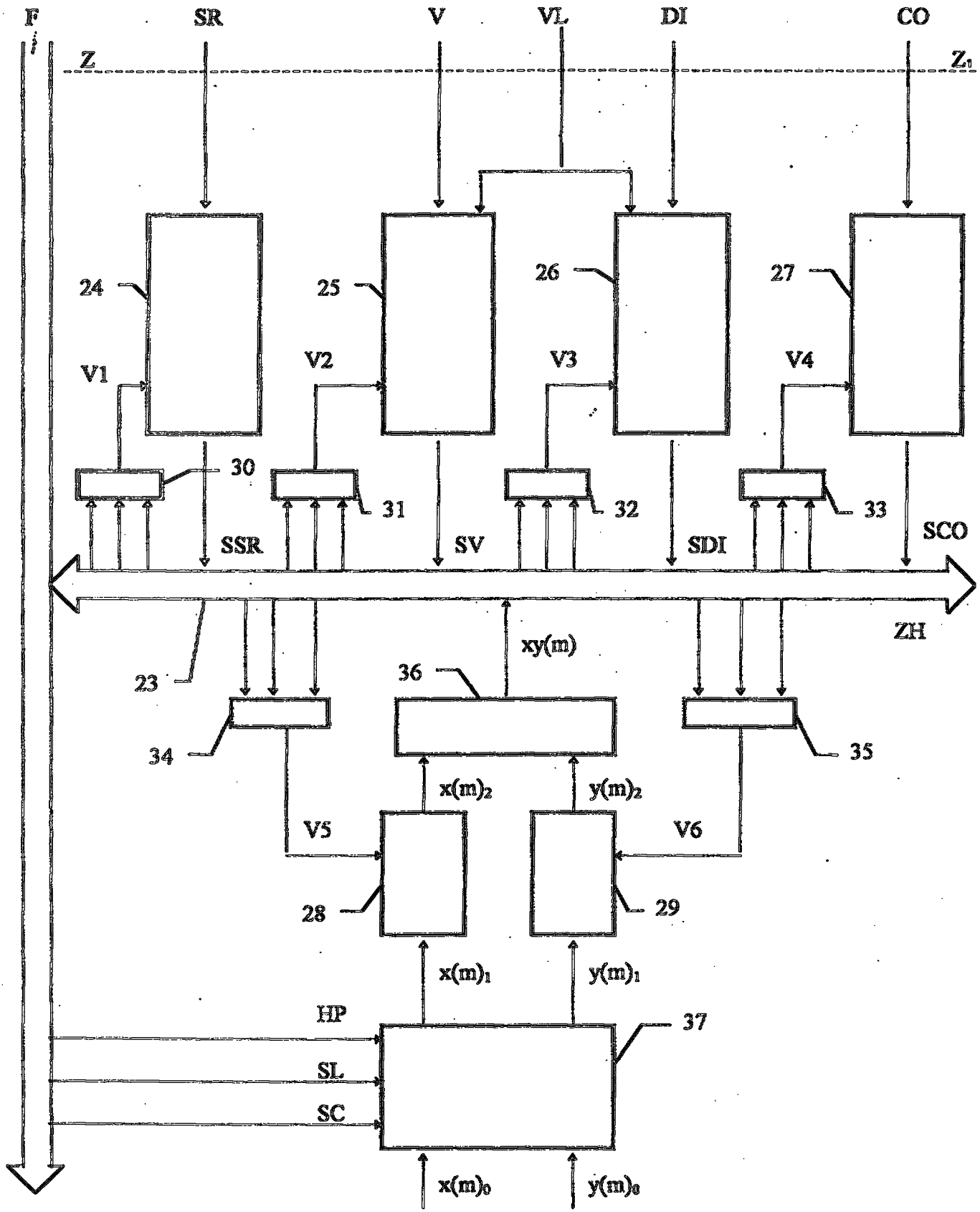

F
SR
V
VL
DI
CO
Z
Z1
24
25
26
27
V1
V2
V3
V4
30
31
32
33
SSR
SV
SDI
SCO
ZH
23
xy(m)
36
34
35
x(m)2
y(m)2
V5
V6
28
29
x(m)1
y(m)1
HP
37
SL
SC
x(m)0
y(m)0

FIG. 11

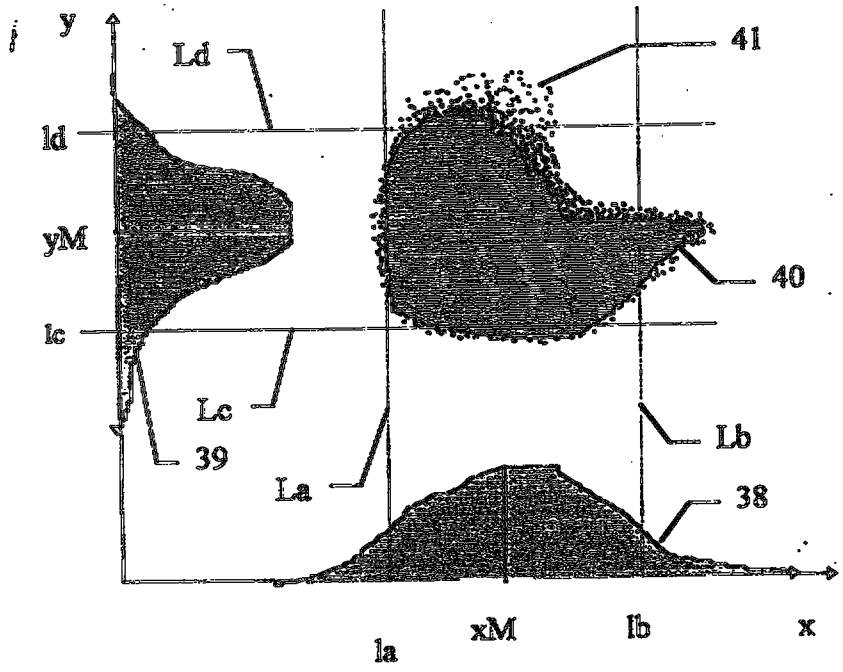
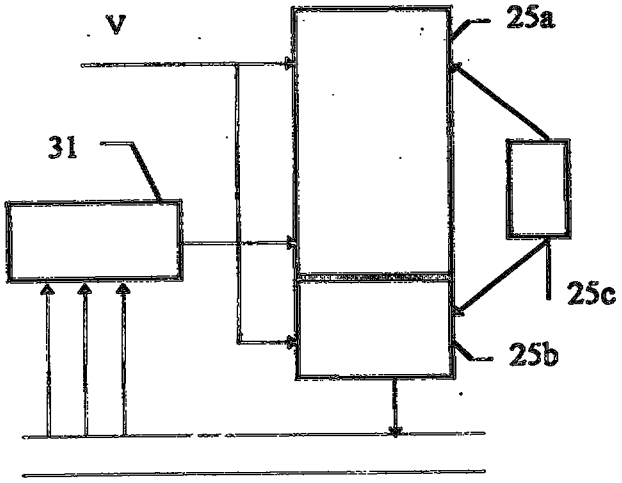


FIG. 12

FIG. 13



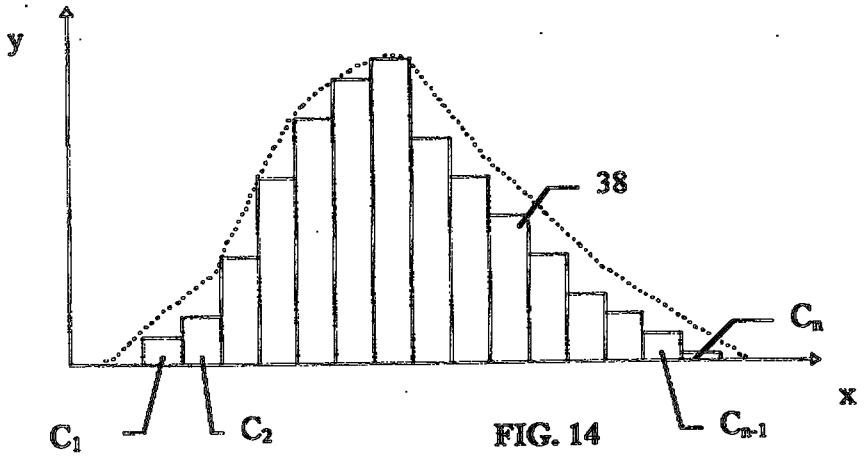
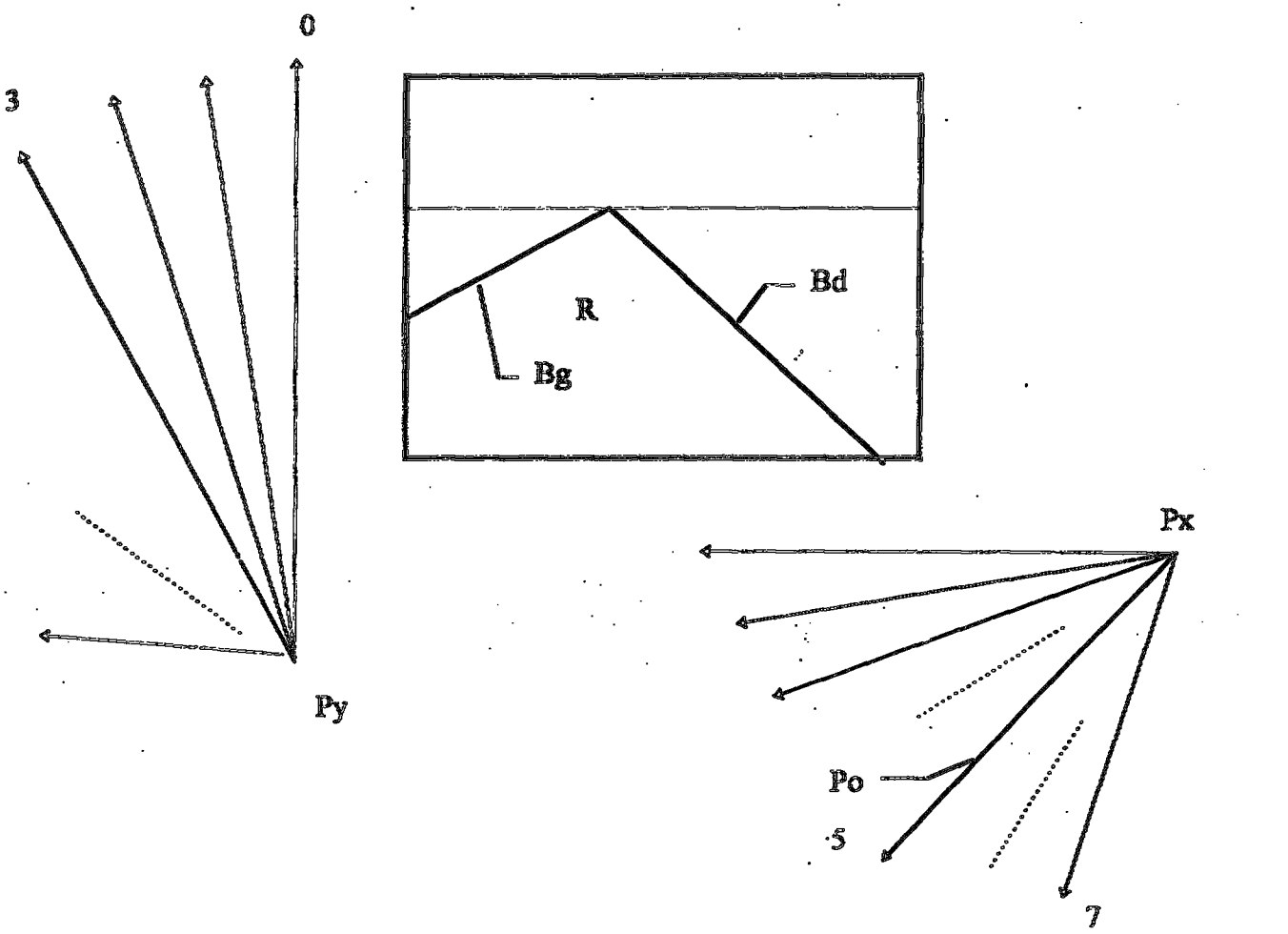


FIG. 14

FIG. 15



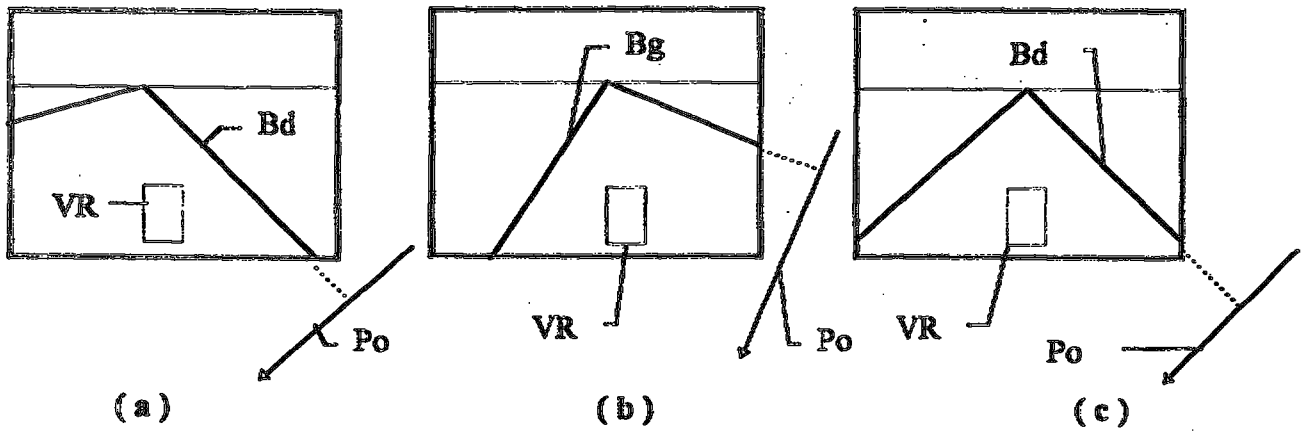
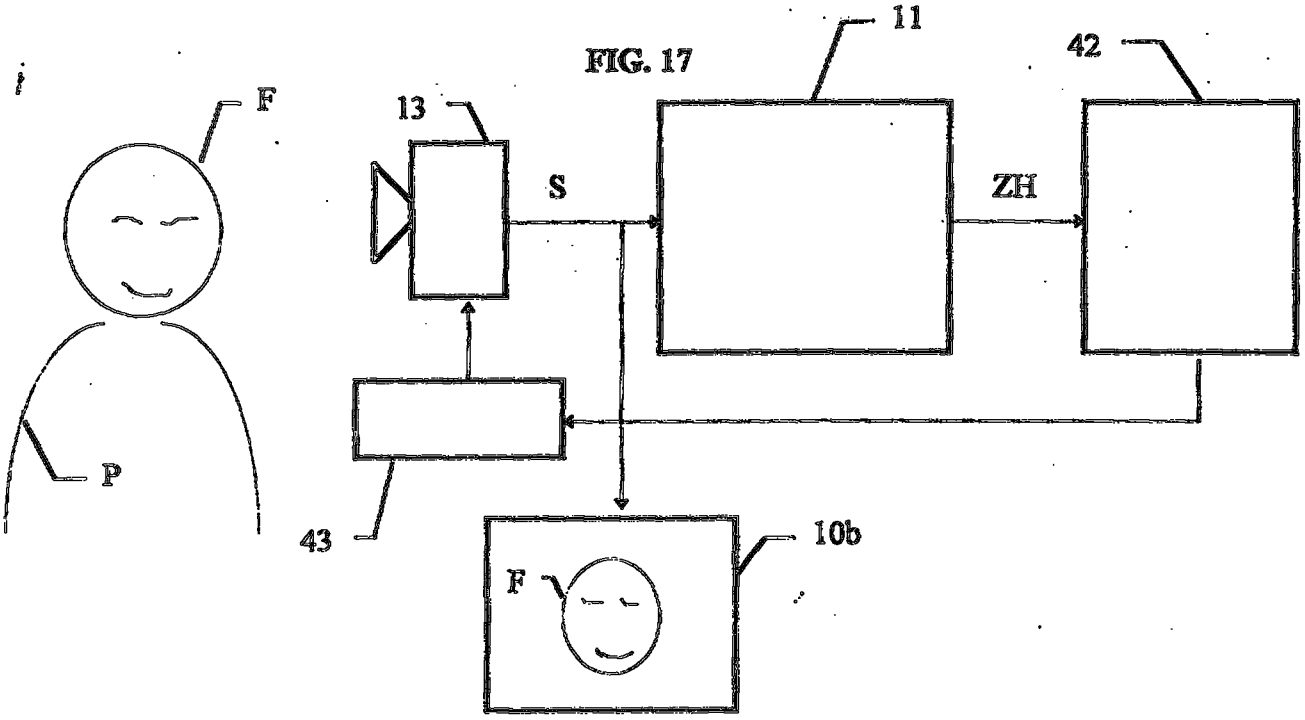


FIG. 16

FIG. 17



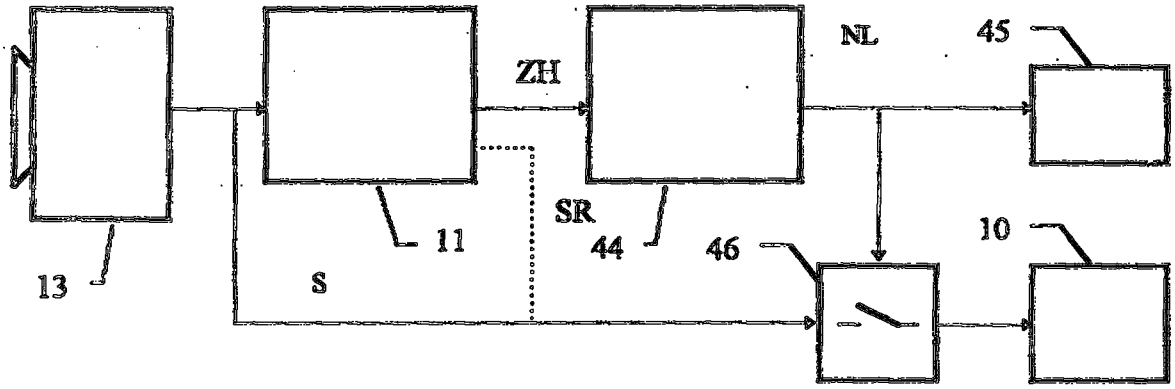


FIG. 18

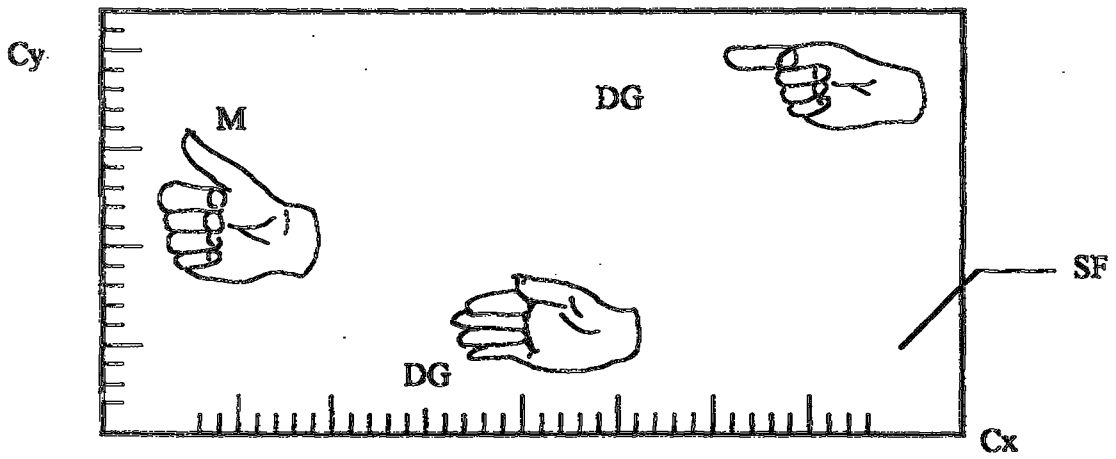


FIG. 19

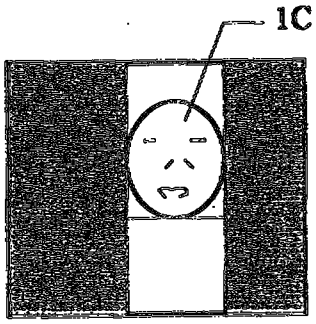


FIG.20

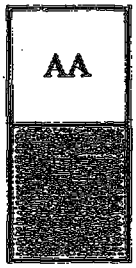


FIG.21

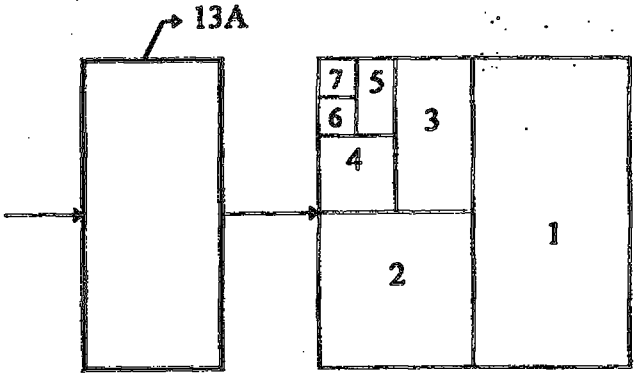


FIG.22

* C L M S *

REVENDICATIONS

1. Procédé, opérant en temps réel, pour le repérage et la localisation d'une zone en mouvement relatif dans une scène observée par un système d'observation à sortie constituée par un signal vidéo numérique du type comportant, d'une manière classique, une succession de trames correspondantes de même nature, chacune composée d'une succession de lignes, composées chacune d'une succession de pixels, ainsi que pour la détermination de la vitesse et de la direction orientée du mouvement de déplacement relatif , ledit procédé

consistant à effectuer sur ledit signal vidéo numérique successivement

- un traitement temporel qui consiste, pour chaque position de pixel, à déterminer la différence entre l'amplitude du signal de pixel dans la trame actuelle et l'amplitude d'un signal de pixel représentatif de ses valeurs dans les trames antérieures, et

- un traitement spatial, qui consiste, pour chaque trame dudit signal vidéo numérique, à répartir, en une matrice à nombre de lignes et nombre de colonnes réduits par rapport respectivement au nombre de lignes et au nombre de pixels par ligne dans ledit signal vidéo numérique, uniquement les valeurs, au même instant d'observation, pour une fraction des pixels d'une trame - fraction qui défile par balayage à travers ladite matrice pendant la durée d'une trame -, des signaux représentatifs des variations des valeurs de pixels et à déduire, à partir de l'ensemble des représentations matricielles relatives à un instant d'observation donné, c'est-à-dire à une trame donnée dudit signal vidéo ou numérique, la localisation d'une zone éventuelle en mouvement relatif et l'estimation de ce mouvement s'il existe,

et étant caractérisée en ce que

- il comporte un double traitement préliminaire consistant en

FEUILLE MODIFIEE

- un lissage adaptatif dudit signal vidéo numérique en mettant en oeuvre une constante de temps numérique dont la valeur numérique, pour chacun des pixels dudit signal vidéo numérique et indépendamment pour chacun d'eux, est modifiée en réponse à la variation ou non variation temporelle de la valeur du pixel, et
- à une mise en mémoire d'une trame entière dudit signal de sortie après lissage, d'une part, et de ladite constante de temps pour chaque position de pixel de ladite trame, d'autre part;

- ledit traitement temporel consiste, pour chaque position de pixel, à déterminer l'existence, d'une part, et l'amplitude, d'autre part, d'une variation significative éventuelle de l'amplitude du signal de pixel entre la trame actuelle et la trame juste antérieure lissée, mise en mémoire, et à générer deux signaux numériques , le premier étant un signal binaire, donc monobit, à deux valeurs possibles dont la première représente l'existence d'une telle variation significative et la seconde l'absence d'une telle variation significative entre les deux dites trames actuelle et juste antérieure lissée, la valeur dudit signal binaire modifiant la valeur mémorisée de ladite constante de temps pour le pixel en cause afin respectivement de la diminuer si ledit signal binaire a ladite première valeur et de l'augmenter si ledit signal binaire a ladite seconde valeur, la diminution ou l'augmentation étant réalisée d'une manière quantifiée, tandis que le second signal numérique, dit d'amplitude, est un signal multibits, à nombre limité de bits, quantifiant l'amplitude, de cette variation; et

- ledit traitement spatial consiste, pour chaque trame dudit signal vidéo de sortie,

- à répartir, en ladite matrice et pendant la durée d'une trame, ledit signal binaire, d'une part, et ledit signal numérique d'amplitude, d'autre part,

à déterminer, dans cette double représentation matricielle instantanée dudit signal binaire et dudit signal numérique d'amplitude pour le même pixel, une zone particularisée, dans laquelle à la fois ledit signal binaire a la valeur recherchée représentant la présence, ou respectivement l'absence, de variation significative et ledit signal numérique d'amplitude varie, ou respectivement ne varie pas, d'une valeur significative entre pixels voisins dans ladite matrice suivant une direction orientée à partir d'un pixel d'origine, et ceci pour une même portion de trame, donc à un même instant d'observation, et

- à engendrer des signaux représentatifs de l'existence et de la localisation de la zone de déplacement relatif éventuel, de la valeur de la vitesse relative inter-trames et de la direction orientée parmi plusieurs directions orientées possibles dans le plan d'observation, de ce déplacement relatif éventuel, dans son environnement, à partir de ladite répartition matricielle instantanée de ces deux signaux, binaire et d'amplitude.

2. Procédé selon la revendication 1, caractérisé en ce qu'il consiste, en outre:

- à former les histogrammes des valeurs des signaux répartis matriciellement, d'une part, et les histogrammes de valeurs des inclinaisons de deux axes, à pente variable, de coordonnées dans un plan, d'autre part,
- à repérer, dans chaque histogramme formé, un domaine de variation significative de la valeur traitée et
- à déduire, de chaque domaine repéré, l'existence et la localisation, ainsi que la vitesse et la direction orientée, d'une zone en mouvement relatif.

3. Procédé selon la revendication 1 ou 2, caractérisé en ce que ladite matrice est une matrice carrée à même nombre impair (2/ + 1) de

lignes et de colonnes, et que l'on considère les matrices emboîtées de 3 x 3, 5 x 5, 7 x 7, ...(2*l* + 1) x (2*l* + 1) éléments centrées sur le centre de cette matrice carrée afin de déterminer la matrice emboîtée de plus petite taille dans laquelle ledit signal numérique varie dans une direction orientée à partir dudit centre, la valeur dudit signal binaire représentant un dépassement du seuil selon cette direction.

4. Procédé selon la revendication 1 ou 2, caractérisé en ce que ladite matrice est une matrice hexagonale et que l'on considère les matrices hexagonales emboîtées de taille croissante centrées sur le centre de cette matrice hexagonale afin de déterminer la matrice emboîtée de plus petite taille dans laquelle ledit signal numérique varie dans une direction orientée à partir dudit sommet, la valeur dudit signal binaire représentant un dépassement du seuil selon cette direction.

5. Procédé selon les revendications 1 ou 2, caractérisé en ce que ladite matrice est une matrice en L renversé à une seule ligne et une seule colonne et en ce que l'on considère les matrices emboîtées de 3 x 3 pixels, 5 x 5 pixels, 7 x 7 pixels....(2*l* + 1) x (2*l* + 1) pixels, pour la ligne et la colonne uniques, afin de déterminer la matrice de plus petite taille dans laquelle le signal varie dans une direction orientée, à savoir la ligne de plus grande pente à quantification constante.

6. Procédé selon l'une quelconque des revendications précédentes, caractérisé en ce que ladite constante de temps est de la forme 2^p, *p* étant un nombre inférieur à 16, pouvant donc être exprimé par pas plus de 4 bits, la diminution ou l'augmentation de la constante de temps étant réalisée par la soustraction ou l'addition d'une unité à *p*.

7. Procédé selon la revendication 6, caractérisé en ce que l'on considère des portions successives décroissantes de trames complètes suivant l'algorithme temps - échelle de Mallat et sélectionne la plus

grande de ces portions, qui donne des indications de déplacement, vitesse et orientation, compatible avec la valeur de *p*.

8. Dispositif, fonctionnant en temps réel, pour le repérage et la localisation d'une zone en mouvement relatif dans une scène observée, ainsi que pour la détermination de la vitesse et de la direction orientée du mouvement de déplacement relatif, par mise en oeuvre du procédé selon la revendication 1,

comportant:

- un système d'observation à sortie constituée par un signal vidéo numérique du type comportant, d'une manière classique, une succession de trames correspondantes de même nature, chacune composée d'une succession de lignes, composées chacune d'une succession de pixels,

- un ensemble de traitement temporel déterminant, pour chaque position de pixel, la différence entre l'amplitude du signal de pixel dans la trame actuelle et l'amplitude d'un signal de pixel représentatif de ses valeurs dans les trames antérieures,

- un ensemble de traitement spatial qui, pour chaque trame dudit signal vidéo numérique, répartit en une matrice à nombre de lignes et nombre de colonnes réduits par rapport respectivement au nombre de lignes et au nombre de pixels par ligne dans ledit signal vidéo numérique, uniquement les valeurs, au même instant d'observation, pour une fraction des pixels d'une trame - fraction qui défile par balayage à travers ladite matrice pendant la durée d'une trame -, des signaux représentatifs des variations des valeurs de pixels, et qui déduit, à partir de l'ensemble des représentations matricielles relatives à un instant d'observation donné, c'est-à-dire à une trame donnée dudit signal vidéo numérique, la localisation d'une zone éventuelle en mouvement relatif et l'estimation de ce mouvement s'il existe,

et étant caractérisée en ce qu'il comporte en combinaison:

- des moyens de lissage adaptatif (15) dudit signal vidéo numérique en mettant en oeuvre une constante de temps numérique (CO) dont la valeur numérique, pour chacun des pixels dudit signal vidéo numérique et indépendamment pour chacun d'eux, est modifiée en réponse à la variation ou non variation temporelle de la valeur du pixel;

- des moyens de mémorisation (16) pour une trame entière dudit signal de sortie après lissage (LI), d'une part, et de ladite constante de temps pour chaque position de pixel de ladite trame, d'autre part;

- dans l'ensemble de traitement temporel (15),

- des moyens de comparaison (15) pour déterminer l'existence, d'une part, et l'amplitude, d'autre part, d'une variation significative éventuelle de l'amplitude du signal de pixel entre la trame actuelle et la trame juste antérieure lissée, mise en mémoire, et pour générer un premier signal numérique (DP) , qui est un signal binaire, donc monobit, à deux valeurs possibles dont la première représente l'existence d'une telle variation significative et la seconde l'absence d'une telle variation significative entre les deux dites trames actuelle et juste antérieure lissée, la valeur dudit signal binaire modifiant la valeur mémorisée de ladite constante de temps pour le pixel en cause afin respectivement de la diminuer si ledit signal binaire a ladite première valeur et de l'augmenter si ledit signal binaire a ladite seconde valeur, la diminution ou l'augmentation étant réalisée d'une manière quantifiée, et

- des moyens de calcul (15c) aptes à générer un second signal numérique (CO), dit d'amplitude, qui est un signal multibits, à nombre limité de bits, quantifiant l'amplitude de cette variation; et

- dans l'ensemble de traitement spatial (11), dont les entrées reçoivent de ladite unité de traitement lesdits signaux numériques, binaire et d'amplitude, successifs pour les pixels d'une même trame:

- des moyens pour répartir, en ladite matrice (21) et pendant la durée d'une trame, ledit signal binaire, d'une part, et ledit signal numérique d'amplitude, d'autre part,
- des moyens pour déterminer, dans cette double représentation matricielle instantanée dudit signal binaire et dudit signal numérique d'amplitude pour le même pixel, une zone particularisée, dans laquelle à la fois ledit signal binaire a la valeur recherchée représentant la présence, ou respectivement l'absence, de variation significative et ledit signal numérique d'amplitude varie, ou respectivement ne varie pas, d'une valeur significative entre pixels voisins dans ladite matrice suivant une direction orientée à partir d'un pixel d'origine, et ceci pour une même portion de trame, donc à un même instant d'observation, et
- des moyens pour engendrer des signaux représentatifs de l'existence et de la localisation de la zone de déplacement relatif éventuel, de la valeur de la vitesse relative inter-trames et de la direction orientée parmi plusieurs directions orientées possibles dans le plan d'observation, de ce déplacement relatif éventuel, dans son environnement, à partir de ladite répartition matricielle instantanée de ces deux signaux, binaire et d'amplitude.

9. Dispositif selon la revendication 8, caractérisé en ce que l'ensemble de traitement spatial (17, 18) comporte des premiers moyens de retard (r) en cascade dont chacun impose un retard égal à l'écart temporel entre deux lignes successives et des seconds moyens de retard (d) en cascade pour chaque ligne imposant chacun un retard égal à l'écart temporel entre deux pixels successifs d'une ligne, les sorties de chacun desdits seconds moyens de retard (d) et l'entrée de la cascade desdits seconds moyens de retard (d) de chaque ligne débitant à un instant donné les valeurs dudit signal binaire et dudit signal numérique d'amplitude, à un même instant, vers lesdits moyens de repérage (17a).

10. Dispositif selon la revendication 8 ou 9, caractérisé en ce qu'il comporte également des moyens (24-29) formant les histogrammes des valeurs de sortie de ladite unité de traitement spatial, ainsi que les histogrammes des inclinaisons de deux axes, à pente variable, de coordonnées dans un plan, des moyens pour repérer, dans chaque histogramme, un domaine de variation significative de la valeur traitée, afin de valider ce domaine en leur sortie et de déduire, pour l'ensemble des histogrammes, des signaux de sortie qui signalent et localisent une zone de la scène observée en mouvement relatif, si elle existe, ainsi que la vitesse et la direction orientée de ce mouvement si ladite zone se déplace effectivement par rapport à son environnement.

11. Dispositif selon la revendication 8, 9 ou 10, caractérisé en ce que lesdits moyens de lissage (15c, 15d) comportent une entrée qui reçoit ledit signal vidéo numérique S (PI) et calculent, pour les pixels successifs d'une trame de ce signal vidéo, un signal lissé (LO), dans lequel les variations temporelles du signal vidéo numérique d'entrée sont diminuées, par mise en oeuvre d'un signal de seuil (N) reçu sur une autre entrée et d'une constante de temps (CO) relative à chaque position de pixel d'une trame, dont la valeur est successivement modifiée afin que le lissage conserve, tout en la réduisant, la tendance de variation du signal vidéo numérique entrant, ces moyens de lissage coopérant avec l'unité de mémoire (16) qui reçoit, stocke et restitue les valeurs actualisées, pour chaque position de pixel d'une trame, du signal lissé et de ladite constante de temps et débitent, sur leurs sorties, au moins la succession, pour chaque position de pixel, des valeurs de la constante de temps actualisée et des valeurs d'un signal binaire de dépassement ou non-dépassement dudit seuil par la valeur absolue de la différence entre la valeur du pixel et sa valeur lissée.

FEUILLE MODIFIEE

12. Dispositif selon l'une quelconque des revendications 8 à 11, caractérisé en ce que ladite unité de traitement spatial (17, 18) réalisant la répartition matricielle, par lignes et par colonnes en nombre réduit, des sorties desdits moyens de lissage, à savoir des valeurs successives de la constante de temps (CO) et dudit signal binaire, comporte des moyens de repérage (17a) pour repérer, dans ladite répartition matricielle, une zone de pixels dans laquelle, à la fois, soit la valeur dudit signal binaire correspond à un dépassement de seuil et ladite constante de temps varie entre pixels voisins d'une valeur significative dans une direction, et pour produire des signaux de sortie indiquant la localisation de ladite zone et la vitesse et la direction orientée du déplacement dans ladite zone, soit la valeur dudit signal binaire correspond à un non-dépassement du seuil et ladite constante de temps ne varie pas entre pixels voisins.

13. Application du dispositif selon la revendication 10, 11 ou 12, pour aider la conduite d'un véhicule terrestre sur une route, respectivement d'un véhicule aérien au voisinage d'une piste d'aéroport, caractérisé en ce que le dispositif comporte en outre des moyens de représentation des bords droit Bd et gauche Bg de la route, respectivement de la piste, et des moyens pour orienter au moins un des axes, à pente variable, de coordonnées pour qu'il se maintienne sensiblement orthogonal au bord correspondant (position Po).

14. Application du dispositif selon l'une quelconque des revendications 1 à 12, pour surveiller l'état de vigilance du conducteur d'un véhicule automobile, afin de détecter une tendance éventuelle à la somnolence de celui-ci, consistant à:

- générer un signal vidéo numérique représentant
  - initialement les images successives de la face du conducteur et

- ensuite, d'une manière continue et en temps réel, les images successives de seulement les yeux du conducteur;

- traiter ledit signal vidéo relatif seulement aux yeux, du conducteur afin de, successivement et en temps réel,

- détecter, dans lesdites images de seulement les yeux, les mouvements verticaux des paupières représentant le clignement de celles-ci,
- déterminer les cadences successives de ces mouvements et
- détecter les cadences qui sont inférieures à un seuil de clignement des paupières qui correspond à la transition entre l'état éveillé et l'état de somnolence du conducteur, et

- déclencher un signal d'alarme apte à alerter le conducteur dès que lesdites cadences franchissent ledit seuil.

REPLACED SHEETS SG
☐ Description : paged ..... to paged .....
☑ Claim(s) : paged 59 to paged 61
☐ Drawing(s) : paged ..... to paged .....
☐ Abstract : paged ..... to paged .....

CLAIMS

1. Process operating in real time for identifying and localizing an area in relative movement in a scene observed by an observation system with an output constituted by a digital video signal of the type comprising a succession of corresponding frames each composed of a succession of lines, each composed of a succession of pixels, and for determining the speed and oriented direction of the displacement, said process being characterized in that it consists in carrying out the following sequence of operations on the digital video output signal:
   - a smoothing processing on said digital video output signal using a digital time constant, the numerical value of which may be modified independently for each pixel of said output signal;
   - a memorization, on the one hand, of a frame of said output signal after smoothing and, on the other hand, of the smoothing time constant associated with said frame;
   - a temporal processing for each pixel position, consisting in determining firstly the existence and secondly the amplitude of a significant variation in the amplitude of the pixel signal between the current frame and the immediately previous smoothed and memorized frame, and generating two digital signals, the first signal being a binary, or single bit, signal with two possible values, one of which represents the existence and the other represents the absence of a significant variation between two successive frames, the value of said binary signal modifying the memorized value of said time constant in order to reduce it if said signal represents a significant variation and in order to increase it if said signal does not represent such a variation, the reduction or increase being done in a quantified

This Sheet Has Been Replaced

manner, whereas the second digital signal, i.e. the amplitude signal, is a multibit signal with a limited number of bits quantifying the amplitude of this variation; and

- a spatial processing consisting of the following steps for each frame of the digital video input signal :
  - distributing only the values of a fraction of the pixels in a frame at a given observation instant (a fraction which is scanned across said matrix throughout the duration of the frame) firstly of said binary signal and secondly of said digital amplitude signal, into a matrix with a number or rows and columns that is small compared with the number of lines and the number of pixels per line in the video signal, respectively, in order to characterize the values of the pixels,
  - determining in this double instantaneous matrix representation, a particular area in which said binary signal has the searched value, representing the presence or absence of a significant variation, and said digital amplitude signal varies, or does not vary, by a significant value for adjacent pixels in the matrix along an oriented direction starting from an origin pixel, within the same portion of the frame and therefore at the same observation instant, and
  - generating signals representing the existence and localization of the area in relative displacement and the relative inter-frame speed and the oriented direction of this displacement if any, relative to its environment, infered from the instantaneous matrix distribution of these two binary and amplitude digital signals.

This Sheet Has Been Replaced

2. Process according to claim 1 characterized in that it also consists of:
   - forming firstly histograms of the signal values distributed in matrices, and secondly histograms of the inclinations of two coordinate axes, with variable slope, in a plane,
   - identifying a domain in which there is a significant variation of the processed values in each histogram formed, and
   - for each identified domain, deducing whether or not there is an area in relative movement, and if so its localization, speed and oriented direction.

3. Process according to claim 1 or 2, characterized in that said matrix is a square matrix with the same odd number of rows and columns $(2I + 1)$, and nested matrices containing 3 x 3, 5 x 5, 7 x 7, ...$(2I + 1)$ x $(2I + 1)$, elements, centered on the center of this square matrix, are considered in order to determine the smallest nested matrix in which said digital signal varies along an oriented direction starting from said center, the value of said binary signal showing that the threshold along this direction has been exceeded.

4. Process according to claim 1 or 2, characterized in that said matrix is a hexagonal matrix, and nested hexagonal matrices of increasing size, centered on the center of this hexagonal matrix, are considered, in order to determine the smallest nested matrix in which said digital signal varies along an oriented direction.

5. Process according to claim 1 or 2, characterized in that said matrix is an inverted L matrix with a single row and a single column, and nested matrices containing 3 x 3 pixels, 5 x 5 pixels, 7 x 7 pixels,.... $(2I + 1)$ x $(2I + 1)$ pixels are considered

This Sheet Has Been Replaced

for the single line and the single column, in order to determine the smallest matrix in which the signal varies along an oriented direction, namely the line with the steepest slope and constant quantification.

6. Process according to any one of the previous claims, characterized in that said time constant is of the form 2^p, where p is a number less than 16, and which can then be expressed in not more than 4 bits, the time constant being reduced or increased by subtracting or adding one unit to p.

7. Process according to claim 6, characterized in that successive decreasing portions of complete frames are considered using the Mallat time-scale algorithm and the largest of these portions, which gives displacement, speed and orientation indications compatible with the value of p, is selected.

8. Device operating in real time for identifying and localizing an area in relative movement in a scene observed by an observation system with an output constituted a digital video signal of the type comprising a succession of corresponding frames, successive lines in each corresponding frame and successive pixels in each line, and for determining the speed and oriented direction of this displacement, by using the process described above, said device receiving said video output signal as input and being characterized in that it comprises in combination:

- means(15) for smoothing said digital video output signal making use of a digital time constant (CO), the numerical value of which may be modified independently for each of said output signal pixels;
- means (16) for memorizing firstly a frame of said output signal after smoothing (LI) and secondly the smoothing time constant (CI) associated with said frame,

This Sheet Has Been Replaced

- a temporal processing unit (15) for analyzing the variations with time of the amplitude of the pixel signal, for the same pixel position, between the current frame and the immediately preceding smoothed and memorized frame of said digital signal, said unit comprising a memory unit (16) capable of receiving, storing and retrieving information about the corresponding previous smoothed frame, together with comparison means (15a) to determine if the absolute value of the difference, between the current pixel signal and a representative value of the pixel signal for the same pixel position in the previous frame stored in said memory unit, exceeds or does not exceed a threshold, for each position of the pixel in the input video signal frame, by generating a binary, or single bit, signal (DP) with two values, one of these two values showing that the threshold is exceeded and the other showing that it is not exceeded, and computing means (15c) capable of determining a multibit digital amplitude signal with a small number of bits, the value of which depends on the amplitude of the variation in the value of the same pixel between the current frame and the immediately preceding, smoothed and memorized, frame of the digital video signal; and
- a spatial processing unit (11), into which the outputs from said successive binary and digital amplitude signals for the pixels in a given frame produced by the temporal processing unit are input, said spatial processing unit comprising means for characterizing the amplitude values of the pixels, these means distributing only said binary and digital amplitude signals relative to the same instant, i.e. to a single frame, into a matrix (21) with fewer rows and columns than the number of lines and the number of pixels per line respectively in a

This Sheet Has Been Replaced

frame of said digital video signal, the frame being scanned across said matrix throughout the duration of a frame, spotting means for determining an area of pixels within said matrix in which, at this instant, the binary signal has the searched value, and means of determining an area of pixels in said matrix, into which, at the same instant, the digital amplitude signal for adjacent pixels varies by a significant quantity, and means which, in response to the indications from the two preceeding means, generate signals representing this area of pixels and therefore the existence and localization of an area with relative movement within the observed scene and the relative inter-frame speed and oriented direction of this area when it actually moves relative to its environment.

9. Device according to claim 7, characterized in that said spatial processing unit includes first delay means (*r*) in cascade, each of which imposes a delay equal to the time difference between two successive lines and second delay means (*d*) in cascade, for each line, each importing a delay equal to the time difference between two successive pixels in a line, the outputs from all said second delay means (*d*) and the input to the cascade of said second delay means (*d*) on each line delivering at a given instant the values of said binary signal and said digital amplitude signal at the same instant, to said identification means (17a).

10. Device according to claim 8 or 9, characterized in that it also comprises means (24-29) for forming histograms of the output values from said spatial processing unit, and histograms of the inclinations of two, variable slope, coordinates axes within a plane, means for identifying a domain, in each histogram, in

This Sheet Has Been Replaced

which there is a significant variation in the processed value, in order to validate this domain at the output, and to deduce, for all histograms, output signals that identify and localize an area of relative movement in the observed scene, if any, and the speed and oriented direction of this movement if said area is effectively moving with respect to its environment.

11. Device according to claim 8, 9 or 10, characterized in that said smoothing means (15c, 15d) include an input receiving said digital video signal S(PI) and derive, for each successive pixel in a frame of said video signal, a smoothed signal (LO) in which the temporal variations of the input digital video signal are reduced by using a threshold signal received on another input and a time constant (CO) related to each pixel position in a frame, the value of which is modified successively so that the smoothing maintains, although reduced, the trend in the variations of the incoming digital video signal, said smoothing means working in cooperation with said memory unit (16) that receives, stores and retrieves the updated values of the smoothed signal and said time constant, for each portion of pixels in a frame, and output, for each pixel position, at least the succession of values of the updated time constant and of the values of the binary signal that shows whether or not said threshold is exceeded by the absolute value of the difference between the value of the pixel and its smoothed value.

12. Device according to any one of claims 8 to 11, characterized in that said spatial processing unit (17, 18) that performs the matrix distribution, into a reduced number of rows and columns, of the outputs from said smoothing means, namely the successive values of the time constant (CO) and said binary signal, comprises identification means (17a) to identify in said matrix distribution, an area of

This Sheet Has Been Replaced

pixels in which simultaneously either the value of said binary signal corresponds to a threshold being exceeded and said time constant varies between adjacent pixels by a significant value in one direction and for producing output signals indicating the localization of said area and the speed and oriented direction of displacement in this area, or the value of said binary signal corresponds to a threshold not being exceeded and said time constant does not vary between adjacent pixels.

13. Application of the device according to claim 10, 11 or 12 to assist in driving a land vehicle on a road, or in flying an aircraft close to an airport runway, characterized in that the device also includes means for representing the right and left sides Bd and Bg of the road or respectively the runway, and means for orienting at least one of the coordinate axes at variable slope, so that it remains approximately orthogonal to the corresponding side (position P_0).

14. Application of any of the preceeding claims to the watching of the vigilance state of the driver of a motor vehicle, for detecting a possible tendency sleepiness thereof, consisting in :

(a) producing a video signal representing

- initially the successive images of the face of the driver and then

- continuously, in real time, the successive images of the eyes of the driver only,

(b) processing said video signal relative only to said eyes for successively, in real time,

- detecting in said image of the eyes, the vertical movements of the eye-lids representing the blinkings thereof,

- determining the rates of these movements and

- detecting the rates which are lower than a threshold blinking rate corresponding substantially to

This Sheet Has Been Replaced

* C L M S *

the transition between the awake state and the sleepiness state of said driver, and

(c) starting a warning signal alerting the driver as soon as said threshold blinking rate is crossed.

This Sheet Has Been Replaced

REPLACEMENT SHEETS SG
☐ Description : paged ... to paged ....
☑ Claim(s) : paged 59 to paged 66.
☐ Drawing(s) : paged ..... to paged ....
☐ Abstract : paged ... to paged ...
DATE OF RECEIPT 13/11/00

REPLACEMENT SHEET
9806372-0
IPO 13 NOV 200

CLAIMS

1. A process, operating in real time, for identifying and localising an area in relative movement in a scene observed by an observation system with an output constituted by a digital video signal comprising, conventionally, a succession of corresponding frames of the same nature, each composed of a succession of lines, each made of a succession of pixels, as well as for determining the velocity and the oriented of the relative displacement movement, whereas the said process consists in performing on the said digital video signal successively

- a temporal processing that consists, for each pixel position, in determining the difference between the amplitude of the pixel signal in the current frame and the amplitude of a pixel signal representative of its values in the previous frames, and

- a spatial processing that consists, for each frame of the said digital video signal, in distributing over a matrix with a smaller number of lines and a smaller number of pixels than respectively the number of lines and the number of pixels per line in the said digital video signal, exclusively the values, at the same observation instant, for a fraction of the pixels of a given frame (a fraction scanned through the said matrix for the duration of a frame), of signals representative of the variations in the pixel values and in deriving, from the set of matrix representations relating to a given observation instant, i.e. a given frame of the said video or digital signal, the localisation of a possible area in relative movement and assessing the said movement if any,

and characterised in that

- it comprises a double preliminary process consisting of

- adaptative smoothing of the said digital video signal while implementing a digital time constant whose digital value for each of the pixels of the said digital video signal and independently for each of them, is modified in response to the temporal variation or non-variation of the pixel value, and

- storing a whole frame of the said output signal after smoothing, on the one hand, and the said time constant for each pixel position in the said frame, on the other hand;

- the said temporal processing consists, for each pixel position, in determining the existence, on the one hand, and the amplitude, on the other, of a possible significant variation in amplitude of the pixel signal between the current frame and the frame smoothed and stored just previously, and in generating two digital signals, whereas the first is a binary signal, hence monobit, with two possible values, the first of which represents the existence of such a significant variation and the second the absence of such a significant variation between both said frames, the current and the one smoothed just previously, whereas the value of the said binary signal modifies the stored value of the said time constant for the pixel in question in order respectively to reduce the said time constant if the said binary signal has the said first value and to increase the time constant if the said binary signal has the second value, whereby the decrease or increase in value is quantified, while the second so-called amplitude, digital signal is a multibit signal, with a limited number of bits, quantifying the amplitude of this variation; and

- the said spatial processing consists, for each frame of the said output video signal,

- in distributing within the said matrix and throughout the duration of a frame, the said binary signal, on the one hand, and the said amplitude digital signal, on the other,

- in determining, in this double instantaneous matrix representation of the said binary signal and of the said amplitude digital signal for the same pixel, in which at the same time the said binary signal has the requested value representing the presence, or respectively the absence, of any significant variation and the said amplitude digital signal varies, or respectively does not vary, by a significant value between neighbouring pixels in the said matrix along an oriented direction from a starting pixel and this for the same portion of frame, hence at the same instant of observation, and

- in generating signals representative of the existence and of the localisation of the area of possible relative displacement, of the value of the relative inter-frame velocity and of the oriented direction among several possible oriented directions in the observation plane, of this possible relative displacement, in its environment, from the said instantaneous matrix distribution of both these signals, binary and amplitude signals.

2. Process according to claim 1 characterized in that it also consists of:

- forming firstly histograms of the signal values distributed in matrices, and secondly histograms of the inclinations of two coordinate axes, with variable slope, in a plane,
- identifying a domain in which there is a significant variation of the processed values in each histogram formed, and
- for each identified domain, deducing whether or not there is an area in relative movement, and if so its localization, speed and oriented direction.

3. Process according to claim 1 or 2, characterized in that said matrix is a square matrix with the same odd number of rows and columns $(2l + 1)$, and nested matrices containing 3 x 3, 5 x 5, 7 x 7, ...$(2l + 1)$ x $(2l + 1)$, elements, centered on the center of this square matrix, are considered in order to determine the smallest nested matrix in which said digital signal varies along an oriented direction starting from said center, the value of said binary signal showing that the threshold along this direction has been exceeded.

4. Process according to claim 1 or 2, characterized in that said matrix is a hexagonal matrix, and nested hexagonal matrices of increasing size, centered on the center of this hexagonal matrix, are considered, in order to determine the smallest nested matrix in which said digital signal varies along an oriented direction.

5. Process according to claim 1 or 2, characterized in that said matrix is an inverted L matrix with a single row and a single column, and nested matrices containing 3 x 3 pixels, 5 x 5 pixels, 7 x 7 pixels,.... $(2l + 1)$ x $(2l + 1)$ pixels are considered for the single line and the single column, in order to determine the

REPLACEMENT SHEET
13 NOV 2000

smallest matrix in which the signal varies along an oriented direction, namely the line with the steepest slope and constant quantification.

6. Process according to any one of the previous claims, characterized in that said time constant is of the form 2^p, where p is a number less than 16, and which can then be expressed in not more than 4 bits, the time constant being reduced or increased by subtracting or adding one unit to p.

7. Process according to claim 6, characterized in that successive decreasing portions of complete frames are considered using the Mallat time-scale algorithm and the largest of these portions, which gives displacement, speed and orientation indications compatible with the value of p, is selected.

8. A device operating in real time, for identifying and localising an area in relative movement in a scene observed by an observation system, as well as for determining the velocity and the oriented of the relative displacement movement, by implementing the process according to claim 1 comprising

- an observation system with an output constituted by a digital video signal consisting of, conventionally, a succession of corresponding frames of the same nature, each composed of a succession of lines, each made of a succession of pixels

- a temporal processing assembly that determines, for each pixel position, the difference between the amplitude of the pixel signal in the current frame and the amplitude of a pixel signal representative of its values in the previous frames,

- a spatial processing assembly that, for each frame of the said digital video signal, distributes over a matrix with a smaller number of lines and a smaller number of pixels than respectively the number of lines and the number of pixels per line in the said digital video signal, exclusively the values, at the same observation instant, for a fraction of the pixels of a given frame (a fraction scanned through the said matrix for the duration of a frame), of signals representative of the variations of pixels and derives, from the set of matrix representations relating to a given observation instant, i.e. a given frame of the

said video or digital signal, the localisation of a possible area in relative movement and assessing the said movement if any,

and characterised in that it comprises in combination

- means for adaptative smoothing (15) of the said digital video signal while implementing a digital time constant (CO) whose digital value for each of the pixels of the said digital video signal and independently for each of them, is modified in response to the temporal variation or non-variation of the pixel value, and

- means for storing (16) a whole frame of the said output signal after smoothing (LI), on the one hand, and the said time constant for each pixel position in the said frame, on the other hand;

- in the said temporal processing assembly (15)

- comparison means (15) for determining the existence, on the one hand, and the amplitude, on the other, of a possible significant variation in amplitude of the pixel signal between the current frame and the frame smoothed and stored just previously, and for generating a first digital signal (DP); which is a binary signal, hence monobit, with two possible values, the first of which represents the existence of such a significant variation and the second the absence of such a significant variation between both said frames, the current and the one smoothed just previously, whereas the value of the said binary signal modifies the stored value of the said time constant for the pixel in question in order respectively to reduce the said time constant if the said binary signal has the said first value and to increase the time constant if the said binary signal has the second value, whereby the decrease or increase in value is quantified, and

- calculation means (15c) capable of generating a second so-called amplitude, digital signal (CO), which is a multibit signal, with a limited number of bits, quantifying the amplitude of this variation; and

- in the spatial processing assembly (11), whose inputs receive from the said processing unit the said successive binary and amplitude digital signals for the pixels of the same frame:

- means for distributing within the said matrix (21) and throughout the duration of a frame, the said binary signal, on the one hand, and the said amplitude digital signal, on the other,
- means for determining, in this double instant matrix representation of the said binary signal and of the said amplitude digital signal for the same pixel, a particularised area, in which at the same instant the said binary signal has the requested value representing the presence, or respectively the absence, of significant variation and the said amplitude digital signal varies, or respectively does not vary, by a significant value between neighbouring pixels in the said matrix along an oriented direction from a starting pixel and this for the same portion of frame, hence at the same instant of observation, and
- means for generating signals representative of the existence and of the localisation of the area of possible relative displacement, of the value of the relative inter-frame velocity and of the oriented direction among several possible oriented directions in the observation plane, of this possible relative displacement, in its environment, from the said instantaneous matrix distribution of both these signals, binary and amplitude signals.

9. Device according to claim 8, characterized in that said spatial processing unit includes first delay means (*r*) in cascade, each of which imposes a delay equal to the time difference between two successive lines and second delay means (*d*) in cascade, for each line, each importing a delay equal to the time difference between two successive pixels in a line, the outputs from all said second delay means (*d*) and the input to the cascade of said second delay means (*d*) on each line delivering at a given instant the values of said binary signal and said digital amplitude signal at the same instant, to said identification means (17a).

10. Device according to claim 8 or 9, characterized in that it also comprises means (24-29) for forming histograms of the output values from said spatial processing unit, and histograms of the inclinations of two, variable slope, coordinates axes within a plane, means for identifying a domain, in each histogram, in which there is a significant variation in the processed value, in

order to validate this domain at the output, and to deduce, for all histograms, output signals that identify and localize an area of relative movement in the observed scene, if any, and the speed and oriented direction of this movement if said area is effectively moving with respect to its environment.

11. Device according to claim 8, 9 or 10, characterized in that said smoothing means (15c, 15d) include an input receiving said digital video signal S(PI) and derive, for each successive pixel in a frame of said video signal, a smoothed signal (LO) in which the temporal variations of the input digital video signal are reduced by using a threshold signal received on another input and a time constant (CO) related to each pixel position in a frame, the value of which is modified successively so that the smoothing maintains, although reduced, the trend in the variations of the incoming digital video signal, said smoothing means working in cooperation with said memory unit (16) that receives, stores and retrieves the updated values of the smoothed signal and said time constant, for each portion of pixels in a frame, and output, for each pixel position, at least the succession of values of the updated time constant and of the values of the binary signal that shows whether or not said threshold is exceeded by the absolute value of the difference between the value of the pixel and its smoothed value.

12. Device according to any one of claims 8 to 11, characterized in that said spatial processing unit (17, 18) that performs the matrix distribution, into a reduced number of rows and columns, of the outputs from said smoothing means, namely the successive values of the time constant (CO) and said binary signal, comprises identification means (17a) to identify in said matrix distribution, an area of pixels in which simultaneously either the value of said binary signal corresponds to a threshold being exceeded and said time constant varies between adjacent pixels by a significant value in one direction and for producing output signals indicating the localization of said area and the speed and oriented direction of displacement in this area, or the value of said binary signal corresponds to a threshold not being exceeded and said time constant does not vary between adjacent pixels.

REPLACEMENT SHEET

13 NOV 2000

13. Application of the device according to claim 10, 11 or 12 to assist in driving a land vehicle on a road, or in flying an aircraft close to an airport runway, characterized in that the device also includes means for representing the right and left sides B_d and B_g of the road or respectively the runway, and means for orienting at least one of the coordinate axes at variable slope, so that it remains approximately orthogonal to the corresponding side (position P_o).

14. Application of the device according to any of claims 8 to 12 for monitoring the vigilance state of the driver of a motor vehicle, for detecting a possible tendency sleepiness thereof, consisting in :

(a) producing a video signal representing

- initially the successive images of the face of the driver and then
- continuously, in real time, the successive images of the eyes of the driver only,

(b) processing said video signal relative only to said eyes for successively, in real time,

- detecting in said image of the eyes, the vertical movements of the eyelids representing the blinkings thereof,
- determining the rates of these movements and
- detecting the rates which are lower than a threshold blinking rate corresponding substantially to the transition between the awake state and the sleepiness state of said driver, and

(c) starting a warning signal alerting the driver as soon as said threshold blinking rate is crossed.

9806372-0
FR 05

PCT ORGANISATION MONDIALE DE LA PROPRIETE INTELLECTUELLE
Bureau international

DEMANDE INTERNATIONALE PUBLIEE EN VERTU DU TRAITE DE COOPERATION EN MATIERE DE BREVETS (PCT)

(51) Classification internationale des brevets 6 : G06T 7/20 — A1

(11) Numéro de publication internationale: WO 98/05002

(43) Date de publication internationale: 5 février 1998 (05.02.98)

(21) Numéro de la demande internationale: PCT/FR97/01354

(22) Date de dépôt international: 22 juillet 1997 (22.07.97)

(30) Données relatives à la priorité:
96/09420 26 juillet 1996 (26.07.96) FR

(71) Déposant *(pour tous les Etats désignés sauf US)*: CARLUS MAGNUS LIMITED [-/-]; Victoria House, Main Street, Gibraltar (GI).

(72) Inventeur; et
(75) Inventeur/Déposant *(US seulement)*: PIRIM, Patrick [FR/FR]; 56, rue Patay, F-75013 Paris (FR).

(74) Mandataire: LE BRUSQUE, Maurice; Cabinet Harle & Phelip, 21, rue de la Rochefoucauld, F-75009 Paris (FR).

(81) Etats désignés: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TR, TT, UA, UG, US, UZ, VN, brevet ARIPO (GH, KE, LS, MW, SD, SZ, UG, ZW), brevet eurasien (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), brevet européen (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), brevet OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

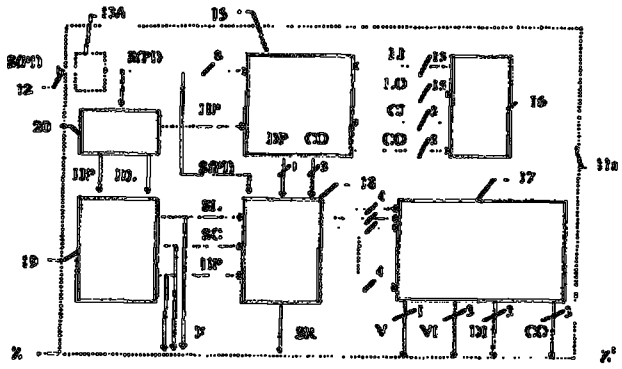
Publiée
Avec rapport de recherche internationale.
Avec revendications modifiées et déclaration.

(54) Title: METHOD AND DEVICE FOR REAL-TIME DETECTION, LOCATION AND DETERMINATION OF THE SPEED AND DIRECTION OF MOVEMENT OF AN AREA OF RELATIVE MOVEMENT IN A SCENE

(54) Titre: PROCEDE ET DISPOSITIF FONCTIONNANT EN TEMPS REEL, POUR LE REPERAGE ET LA LOCALISATION D'UNE ZONE EN MOUVEMENT RELATIF DANS UNE SCENE, AINSI QUE POUR LA DETERMINATION DE LA VITESSE ET DE LA DIRECTION DU DEPLACEMENT

(57) Abstract

A method and device for real-time detection, location and determination of the speed and direction of movement of an area of relative movement in a scene, are disclosed. According to the method, the digital video input signal S(PI) is subjected to a time-based processing step wherein changes in the value of each pixel between one frame and the corresponding previous frame are used to generate a binary signal DP representing a significant change or the lack thereof, and a digital signal CO representing the degree of change; and to a spatial processing step wherein both signals are distributed over a matrix for a single frame passing therethrough, and the relative movement to be sensed as well as the parameters thereof are deduced from the resulting matrix distribution. For this purpose, the device comprises a time processing unit (15) combined with a memory (16) and a spatial processing unit (17) combined with a delay unit (18). Clock (20) and control (19) units are provided for clocking the operation of units (15) and (17).



(57) Abrégé

L'invention a pour objet un procédé et un dispositif, fonctionnant en temps réel, pour le repérage, la localisation, la détermination de la vitesse et de la direction du déplacement en temps réel d'une zone en mouvement relatif dans une scène. Le procédé réalise un traitement temporel du signal vidéo numérique d'entrée S(PI), consistant à déduire, des variations de la valeur de chaque pixel entre une trame et la trame correspondante antérieure, un signal binaire DP de variation ou non-variation significative et un signal numérique CO représentatif de l'importance de cette variation, et un traitement spatial, consistant à répartir sur une matrice par roulement ces deux signaux pour une même trame qui défile à travers la matrice et à déduire de cette répartition matricielle le mouvement relatif recherché et ses paramètres. A cet effet, le dispositif comporte une unité de traitement temporel (15) associée à une mémoire (16) et une unité de traitement spatial (17) associée à une unité à retards (18); les unités d'horloge (20) et de commande (19) cadençant le fonctionnement des unités (15 et 17).