

System zur dynamischen Kontrastmaximierung zwischen Vordergrund und Hintergrund in Bildern oder/und Bildsequenzen

Die Erfindung betrifft ein System zur dynamischen Kontrastmaximierung zwischen Vordergrund und Hintergrund in Bildern oder/und Bildsequenzen.

Für die nahtlose Integration virtueller Objekte in reelle Szenen (auch Augmented Reality genannt) ist es bekannt, dass zwei Probleme zu lösen sind. Zum einen ist die Transformation zwischen Kamerakoordinatensystem und Szenenkoordinatensystem zu bestimmen und zum anderen sind Verdeckungen virtueller Objekte, die in der Regel im Hintergrund abgebildet sind, durch reelle Objekte, die in der Regel im Vordergrund sind, zu bestimmen.

10

Im Anwendungsgebiet Fernsehübertragung bzw. Broadcasting ist man daran interessiert, die beiden genannten Probleme möglichst nicht-mechanisch zu lösen, d.h., ohne den zusätzlichen Einsatz von Markern oder Sensoren. Dies erlaubt eine kürzere Aufbauzeit und reduziert die Komplexität eines insgesamt aufzubauenden Übertragungssystems vor Ort. Gerade bei Live-Übertragungen müssen der Großteil der Geräte zum Einsatzort transportiert und anschließend auf- und abgebaut werden.

Zur Lösung des Transformationsproblems gibt es einerseits hardware-basierte Lösungen wie beispielsweise „encoded heads“ (z.B. von vinten.com), und andererseits software-basierte, visuell arbeitende Lösungen. Beide Lösungen stellen eine Abbildung $x=M(X)$ bereit, die es erlaubt, eine 3D-Koordinate X eines virtuellen Objektes im 3D-Raum in eine Kamerakoordinate x im aktuellen Kamerabild abzubilden.

25

Bisher bekannte Segmentierungs-Methoden benötigen sogenannte Keying-Hinweise, wie z.B. eine bestimmte Hintergrundfarbe (z.B. bei Chroma Keying oder Infrarot-Keying). Diese mechanischen Lösungen haben jedoch die oben genannten Nachteile, da die Keying-Hinweise am Übertragungsort installiert werden müssen.

30

Aufgabe der Erfindung ist es, die obigen Nachteile zu vermeiden.

Zur Lösung dieser Aufgabe wird ein System zur dynamischen Kontrastmaximierung zwischen Vordergrund und Hintergrund in Bildern oder/und Bildsequenzen vorgeschlagen, umfassend wenigstens eine Bildaufnahmeeinrichtung,

35

wenigstens ein mit der Bildaufnahmeeinrichtung verbundenes Analysemodul, das dazu eingerichtet ist, eine Transformation zwischen einem zweidimensionalen Koordinatensystem der Bildaufnahmeeinrichtung und einem zweidimensionalen Koordinatensystem einer Displayvorrichtung zu bestimmen, eine Segmentierung durchzuführen, um in einem Bild wenigstens Verdeckungen von Hintergrundobjekten durch Vordergrundobjekte zu bestimmen,

wenigstens ein mit dem Analysemodul verbundenes Kontrollmodul, wenigstens eine Displayvorrichtung, die dazu eingerichtet ist, Anzeigeeinformation, insbesondere Videoclips, darzustellen, wobei die Bildaufnahmeeinrichtung dazu eingerichtet ist, die auf der Displayvorrichtung dargestellte Anzeigeeinformation aufzunehmen,

wobei das Analysemodul weiter dazu eingerichtet ist, Qualitätsdaten über die Segmentierung zu ermitteln und an das Kontrollmodul zu senden, und wobei das Kontrollmodul dazu eingerichtet ist, die auf der Displayvorrichtung dargestellten Anzeigeeinformationen basierend auf den Qualitätsdaten anzupassen.

Durch die Berücksichtigung von Qualitätsdaten über die Segmentierung ist es möglich, eine Aussage darüber zu erhalten, ob Vordergrund und Hintergrund in ausreichendem Kontrast zueinander sind. Bei einem nicht ausreichenden Kontrast, kann die Anzeigeeinformation angepasst werden, so dass bei unverändertem Vordergrund, ein kontrastreicher Hintergrund erzeugt werden kann.

Das System kann ferner eine Datenbank aufweisen, die mit dem Kontrollmodul und der Displayvorrichtung verbunden ist, wobei die Datenbank dazu eingerichtet ist, eine Abfolge von Anzeigeeinformationen zu speichern, die nacheinander auf der Displayvorrichtung darzustellen sind. Die Datenbank kann beispielsweise Teil eines Medienservers sein. Der Medienserver überträgt abzuspielende Anzeigeeinformationen, beispielsweise Videoclips, an die Displayvorrichtung.

Hierbei ist es bevorzugt, dass das Kontrollmodul dazu eingerichtet ist, die gespeicherte Abfolge der Anzeigeeinformationen zu verändern. In anderen Worten kann das Kontrollmodul den Medienserver anweisen, eine bestimmte Anzeigeeinformation an der Displayvorrichtung auszugeben bzw., darzustellen. Um eine optimale Regelung mittels des Kontrollmoduls zu erreichen, kann die vom Medienserver an der Displayvorrichtung ausgegebene Anzeigeeinformation (Videoclip) auch an das Kontrollmodul übermittelt werden.

Das Analysemodul kann dazu eingerichtet sein, bei der Segmentierung eine Energiefunktion anzuwenden und zu minimieren, wobei die Energiefunktion auf Wahr-

scheinlichkeitsverteilungen von Vordergrundobjekten und Hintergrundobjekten oder/und auf pixelweisen Wahrscheinlichkeitsverteilungen basiert. Die Segmentierung in Vorder- und Hintergrund (wobei auch Zwischenwerte wie Halbverdeckungen oder Unschärfen möglich sind) wird mit Y bezeichnet.

5

Hier kommen Segmentierungs-Methoden zum Einsatz, die keine Keying-Hinweise benötigen, sondern beliebige Vorder- und Hintergrundobjekte trennen können. Dabei wird eine Energiefunktion $E(Y)$ definiert, die minimiert werden soll. $E(Y)$ kann dem negativen Logarithmus der Wahrscheinlichkeit entsprechen, dass eine Verdeckung korrekt bestimmt wurde. Methoden der konvexen Optimierung erlauben es, die Energie effizient und optimal unter Berücksichtigung jedes einzelnen Bildpunktes zu minimieren, d.h., die entsprechende Wahrscheinlichkeit zu maximieren. Für die Minimierung der Energiefunktion $E(Y)$ ist allerdings a-priori Wissen über das Aussehen des Hintergrundes und des Vordergrundes nötig, insbesondere in Form von Wahrscheinlichkeitsverteilungen F (Vordergrund) und B (Hintergrund) bzw. pixelweisen Verteilungen f_i und b_i , wobei i einen bestimmten Bildpunkt/Pixel bezeichnet.

Dieses a-priori Wissen muss aufgrund möglicher Helligkeitsänderungen oder Änderungen des Aussehens aktualisiert werden, was Drifting (das fortschreitende Abweichen eines Modells von der Realität) verursacht. Durch den oben beschriebene Rückkopplungsmechanismus, bei dem Qualitätsdaten über die Segmentierung an das Kontrollmodul übertragen werden und das Kontrollmodul die Anzeigeeinformationen basierend auf den Qualitätsdaten anpasst, kann dem Drift-Problem entgegengewirkt werden. Dies erlaubt deswegen die Verwendung aktueller Segmentierungsmethoden mit verbesserter Leistung. Die Rückkopplung findet insbesondere über die Displayvorrichtung statt, vorzugsweise ein LED-System, wie es z.B. für Bannerwerbung bei vielen Sportereignissen vorhanden ist.

Das Kontrollmodul kann weiter dazu eingerichtet sein, für verbleibende Anzeigeeinformationen der Abfolge zu bestimmen, bei welcher Anzeigeeinformation sich eine erwartete Wahrscheinlichkeitsverteilung des Vordergrunds und eine erwartete Wahrscheinlichkeitsverteilung des Hintergrund der Bildaufnahmeeinrichtung am meisten unterscheiden.

35

Das Kontrollmodul kann auch dazu eingerichtet sein, für eine betreffende Bildaufnahmeeinrichtung eine Wahrscheinlichkeitsabbildung zu berechnen, wobei die Wahrscheinlichkeitsabbildung beschreibt, welche Wahrscheinlichkeitsverteilung

in einem Bild der betreffenden Bildaufnahmeeinrichtung vorliegt, wenn die Displayvorrichtung eine der verbleibenden Anzeigeinformationen darstellt.

Das Kontrollmodul kann ferner dazu eingerichtet sein, die Wahrscheinlichkeitsabbildung basierend auf der zeitveränderlichen 2D-2D-Abbildung $x=F(d)$ zwischen d im Koordinatensystem der Displayvorrichtung und x im Koordinatensystem der Bildaufnahmeeinrichtung zu ermitteln. Diese 2D-2D-Abbildung F kann beispielsweise durch Anwenden von $x=M(X)$ bestimmt werden, wobei X die physische Position X des Display-Punktes d ist.

10

Das Kontrollmodul kann dazu eingerichtet sein, die Wahrscheinlichkeitsabbildung basierend auf einer radiometrischen Kalibrierung zwischen der Displayvorrichtung und der betreffenden Bildaufnahmeeinrichtung zu ermitteln.

15 An der Displayvorrichtung können mehrere beliebig gestaltete Kalibriermarken angeordnet sein, die von der Bildaufnahmeeinrichtung erfassbar sind, um die 2D-2D-Transformation F zwischen dem Koordinatensystem der Displayvorrichtung und dem Koordinatensystem der Bildaufnahmeeinrichtung zu erleichtern. Damit derartige Kalibriermarken nicht regelmäßig auf- und abgebaut werden müssen, können die Displayvorrichtungen solche Kalibriermarken dauerhaft aufweisen. Die Kalibriermarken können auch derartig gestaltet sein, dass eine radiometrische Kalibrierung zwischen Aufnahmevorrichtung und Display möglich ist.

Das Kontrollmodul kann weiter dazu eingerichtet sein, in die Abfolge von Anzeigeeinformationen Kalibrierinformation einzufügen, wobei die Displayvorrichtung dazu eingerichtet ist, die Kalibrierinformation darzustellen. Dabei kann die Kalibrierinformation wenigstens ein einzelnes einfarbiges Bild umfassen, wobei das Bild eine Darstelldauer von wenigstens einem oder genau einem Frame aufweist, und wobei das Bild vorzugsweise eine Farbe aufweist, die einen hohen Kontrast zu einem Vordergrundobjekt aufweist. Wird eine Kalibrierinformation als einfarbiges Bild lediglich für die Dauer eines Frames auf der Displayvorrichtung angezeigt, ist die Kalibrierinformation für einen Betrachter der Displayvorrichtung von bloßem Auge nicht erkennbar. Allerdings wird die an der Displayvorrichtung dargestellte Kalibrierinformation von der Bildaufnahmeeinrichtung aufgenommen, so dass der betreffende Frame dazu verwendet werden kann, Wissen über Vordergrundobjekte und Hintergrundobjekte zu aktualisieren, weil für diesen Frame bekannt ist, welche Farbe der Hintergrund aufweist.

35

Die oben genannte Aufgabe wird auch durch ein Verfahren zur dynamischen Kontrastmaximierung zwischen Vordergrund und Hintergrund in Bildern oder/und Bildsequenzen gelöst, umfassend die Schritte:

- 5 Erfassen von wenigstens einem Bild mittels wenigstens einer Bildaufnahmeeinrichtung,
- Berechnen einer Transformation zwischen einem zweidimensionalen Koordinatensystem der Bildaufnahmeeinrichtung und einem zweidimensionalen Koordinatensystem der Displayvorrichtung,
- 10 Durchführen einer Segmentierung, um in einem Bild wenigstens Verdeckungen von Hintergrundobjekten durch Vordergrundobjekte zu bestimmen,
- Darstellen von Anzeigeeinformation, insbesondere Videoclips, auf einer Displayvorrichtung, wobei die auf der Displayvorrichtung dargestellte Anzeigeeinformation von der Bildaufnahmeeinrichtung aufgenommen wird,
- 15 wobei bei der Segmentierung Qualitätsdaten über die Segmentierung ermittelt werden und die auf der Displayvorrichtung dargestellten Anzeigeeinformationen basierend auf den Qualitätsdaten angepasst werden.

20 Dieses Verfahren kann weitere optionale Schritte umfassen entsprechend den oben für das System beschriebenen Merkmalen. Entsprechend können weitere optionale Aspekte des Verfahrens direkt aus diesen Merkmalen abgeleitet werden.

Nachfolgend wird die Erfindung unter Bezugnahme auf die anliegenden Figuren beispielhaft und nicht einschränkend beschrieben.

25 Fig. 1 zeigt ein schematisches Diagramm des hier vorgestellten Systems.

Ein System 10 zur dynamischen Kontrastmaximierung umfasst mehrere Bildaufnahmeeinrichtungen C1, C2, Ck (durch drei schwarze Punkte unterhalb von C2 angedeutet) und mehrere Analysemodule A1, A2 bis Ak (durch drei schwarze Punkte unterhalb von A2 angedeutet). Jeder Kamera C1, C2 ist ein Analysemodul A1, A2 zugeordnet.

30

Die Kameras C1, C2, Ck leiten geeignete Daten S1k an das jeweilige Analysemodul A1, A2, Ak weiter. Die Analysemodule A1, A2, Ak bestimmen (1) die aktuelle 2D-2D-Abbildung F1, F2, Fk zwischen einer 2D-Koordinate im Displaykoordinatensystem von L in das Bildkoordinatensystem der jeweiligen Kamera C1, C2, Ck und (2) die bildpunktweise Verdeckung. Die Daten S1k sind je nach Analysemodul rein optisch, können jedoch auch Sensordaten von Schrittsensoren, Beschleunigungs-

35

gungssensoren etc. beinhalten. Das Mapping Fk kann manuell kalibriert worden sein, oder mit einem automatischen Kalibrierverfahren.

5 Die Ausgangsdaten S2k der Analysemodule A1, A2, Ak beschreiben die bestimmten Mappings Fk sowie Metadaten über die Verdeckung. Diese Ausgangsdaten werden an ein Kontrollmodul K übertragen.

10 In die andere Richtung (angedeutet durch die Pfeile TCP) vom Kontrollmodul K zu den Analysemodulen Ak fließen Metainformationen zur Verbesserung der Segmentierung.

15 Ferner ist das Kontrollmodul K über ein Netzwerk verbunden mit einer Datenbank, D, in der Anzeigeinformationen (Videoclip) gespeichert sind, die auf dem LED-System L dargestellt werden können. Das Kontrollmodul hat insbesondere Kontroll-Zugriff über das Netzwerk auf einen Medienserver mit einer Videoclip-Datenbank D. Der Medienserver sendet abzuspielende Videoclips bzw. Anzeigeinformationen Z an das LED-System L.

20 Das Kontrollmodul K ist dazu eingerichtet, den Medienserver anzuweisen, bestimmte Anzeigeinformationen bzw. Videoclips Z über ein Bildtransportprotokoll BP, z.B. über HDMI, SDI, Displayport oder SDI over TCP, an die Displayvorrichtung L auszugeben. Zur Regelung wird das aktuell ausgegebene Videobild wiederum an K über BP zurückgegeben.

25 Zu beachten ist insbesondere der Aufbau, mit dem das Kontrollmodul K das Videomaterial (Anzeigeinformation) auswählt, dass auf der Displayvorrichtung L angezeigt werden soll, sowie der Regelkreis mit Feedback zur Kamera C1, C2. Üblicherweise ist das Videomaterial (Videoclips) a priori in einer Playlist organisiert, die während einer Übertragung eines Ereignisses, wie beispielsweise eines Spiels
30 in einer bestimmten Sportart, ohne Möglichkeit auf Adaption abgespielt wird.

Vorliegende ist das Kontrollmodul K dazu eingerichtet, eine die Reihenfolge R von verbleibenden Anzeigeinformationen bzw. Videoclips Z dynamisch anzupassen, um
35 möglichst gut Segmentierungsergebnisse zu erhalten.

Hierzu wird in den Metadaten S2k eine Metrik über die Qualität (Qualitätsdaten) der Segmentierung übertragen, beispielsweise die nach der Minimierung resultierende Segmentierungsenergie $E_k(L)$. Überschreitet diese Energie ein akzeptables Maß in einer oder mehreren Kameras C1, C2, Ck, wird die Videoclip-Reihenfolge

der verbleibenden Abfolge von Videoclips dynamisch angepasst. Wie bereits einleitend erwähnt kommen Segmentierungs-Methoden zum Einsatz, die keine Keying-Hinweise benötigen, sondern beliebige Vorder- und Hintergrundobjekte trennen können. Dabei wird eine Energiefunktion $E(Y)$ definiert, die minimiert werden soll.

- 5 $E(Y)$ kann dem negativen Logarithmus der Wahrscheinlichkeit entsprechen, dass eine Verdeckung korrekt bestimmt wurde. Methoden der konvexen Optimierung erlauben es, die Energie effizient und optimal unter Berücksichtigung jedes einzelnen Bildpunktes zu minimieren, d.h., die entsprechende Wahrscheinlichkeit zu maximieren. Für die Minimierung der Energiefunktion $E(Y)$ ist allerdings a-priori
10 Wissen über das Aussehen des Hintergrundes und des Vordergrundes nötig, insbesondere in Form von Wahrscheinlichkeitsverteilungen F (Vordergrund) und B (Hintergrund) bzw. pixelweisen Verteilungen f_i und b_i .

- In der nachfolgenden Gleichung bezeichnet l das Analysemodul, bei dem die verbleibende Energie $E_k(Y_k)$ am größten ist, d.h., die Segmentierung am wahrscheinlichsten nicht korrekt bzw. am unsichersten ist. Dies ist ein Hinweis auf einen schlechten Kontrast zwischen Vordergründ und Hintergrund.

$$l = \arg_k \max E_k(Y_k)$$

20

- Entsprechend wird beispielsweise für die ausgewählte Kamera Cl die Abbildung $G_l(Z)$ ermittelt, die beschreibt, welche Wahrscheinlichkeitsverteilung (an welchem Pixel) im Kamerabild von Cl vorliegt, wenn das LED-System L die Anzeigeeinformation bzw. den Videoclip Z abspielt. G_l berücksichtigt das 2D-2D-Mapping F_l vom
25 LED-System L zur Kamera Cl und Hintergrundinformationen einer radiometrischen Kalibrierung vom LED-System L zur Kamera Cl .

- Für jeden möglichen in der Abfolge verbleibenden Videoclip Z_m (Anzeigeeinformation) wird die Abbildung $G_l(Z_m)$ bestimmt. Es wird derjenige Clip Z_n als nächstes
30 abgespielt, bei dem die Wahrscheinlichkeitsverteilung F_l des Vordergrundes von Kamera Cl und die erwartete Wahrscheinlichkeitsverteilung G_l des Hintergrundes von Kamera Cl sich am meisten unterscheiden:

$$n = \arg_m \max d(F_l, G_l(Z_m))$$

35

Eine Metrik d zum Vergleich von Vorder- und Hintergrundverteilung ist beispielsweise die Kullback-Leibler-Divergenz.

Das hier beschriebene System 10 und zugehörige Verfahrensschritte ermöglichen es, periodisch, d.h., wenn die verbleibende Energie E_k aufgrund ungenauem Wissens ansteigt, in jedem Kamerabild Wissen über Vorder- und Hintergrundausssehen zu aktualisieren und damit Drifting (das fortschreitende Abweichen eines Modells von der Realität) aktiv zu verhindern.

Die hier beschriebene Displayvorrichtung L kann ein LED-System sein, etwa in der Form von Werbebanden. LED-Systeme L weisen intern eine Ansteuerbox auf, die geeignete vorbereitete Videosignale über ein Bildtransportprotokoll, z. B. HDMI o.ä. entgegennimmt. Ferner weisen LED-Systeme eine Vielzahl von sogenannten Cabinets auf, die jeweils einen kleinen Teil des Gesamtbildes darstellen können und mit der Ansteuerbox verbunden sind. Die Ansteuerbox gibt dabei jedem Cabinet einen Bildausschnitt des ursprünglichen Videos weiter. Es kann zur Vereinfachung angenommen werden, dass ein LED-System einem regulären Bildschirm entspricht, nur dass der Aufbau eines LED-Systems deutlich flexibler gestaltet sein kann, z.B., dass kein rechteckiges Layout gewählt werden muss.

Es wird darauf hingewiesen, dass das hier beschriebene System 10 auch Teil eines übergeordneten Fernsehübertragungssystems sein kann, wie es in der von der gleichen Anmelderin am gleichen Tag eingereichten Anmeldung mit dem Titel „Fernsehübertragungssystem zur Erzeugung angereicherter Bilder“ beschrieben ist. Das hier beschriebene Kontrollmodul kann beim dort beschriebenen System im Steuermodul 18 enthalten sein. Die hier beschriebene Displayvorrichtung kann beim dort beschriebenen LED-System 20 enthalten sein.

25