METHOD AND SYSTEM FOR CROWD LEVEL ESTIMATION

TECHNICAL FIELD

[0001] The present invention generally relates to methods and systems for crowd estimation, and more particularly relates to methods and systems for estimating a crowd size or crowd level.

BACKGROUND OF THE DISCLOSURE

[0002] Many security systems observe areas where crowds form and it is necessary in many circumstances to estimate the size of the crowd for monitoring foot-traffic through the area or for providing services to the area to accommodate the crowd or for other reasons. There are numerous techniques that estimate crowd levels to discover the number of humans in the crowd. The techniques range from simple pixel level techniques such as background subtraction based blob counting to complex pattern recognition techniques such as body part detection and combined head pattern generation. Each technique has its own advantages and disadvantages.

[0003] A single crowd estimation technique may not be suitable for all environmental and crowd conditions. For example, background subtraction techniques have inferior performance when there is an overlap of humans (i.e., an occlusion). Similarly, body part recognition is also affected in cases of occlusions at high crowd densities, thereby reducing the accuracy of the technique. However, combined head pattern techniques are observed to perform better at high crowd densities due to the underlying concept of learning combined head patterns, yet they tend to have lower accuracies at sparse crowd levels or low crowd densities.

[0004] Thus, what is needed is a method and system for real-time crowd level estimation which provides improved accuracy in a variety of crowd conditions and crowd locations. Furthermore, other desirable features and characteristics will become apparent from the subsequent detailed description and the appended claims, taken in conjunction with the accompanying drawings and this background of the disclosure.

SUMMARY

[0005] According to at least one embodiment of the present invention, a method for crowd level estimation is provided. The method includes receiving an input image of a crowd and estimating a crowd level of the crowd in the input image in response to a crowd density level.

[0006] According to another embodiment of the present invention, a system for crowd level estimation is provided. The system includes an input module and a crowd level estimation module. The input module receives an input image of a crowd and the crowd level estimation module estimates a crowd level of the crowd in the input image in response to a crowd density level.

[0007] And according to yet another embodiment, a computer readable medium is provided. The computer readable medium has instructions stored thereon which when compiled onto a processor perform the method for crowd estimation. The method includes receiving an input image of a crowd and determining a region of interest within the input image of the crowd. The method further includes determining a similarity of the region of interest within the input image of the crowd to each of a plurality of models of crowd level spatial variations, each of the plurality of models of crowd level spatial variations generated by modeling spatial variations of each of a

plurality of crowd levels in response to pixel density variations thereof. And finally, the method includes estimating the crowd level of the crowd in the region of interest within the input image in response to a most similar one of the plurality of models of crowd level spatial variations

BRIEF DESCRIPTION OF THE DRAWINGS

[0008] The accompanying figures, where like reference numerals refer to identical or functionally similar elements throughout the separate views and which together with the detailed description below are incorporated in and form part of the specification, serve to illustrate various embodiments and to explain various principles and advantages in accordance with a present embodiment.

[0009] FIG. 1, comprising FIGs. 1A and 1B, depicts illustrations of front end crowd level estimation in accordance with a present embodiment, wherein FIG. 1A depicts a camera arranged to capture images of a crowd and FIG. 1 B depicts images of the crowd captured by the camera.

[0010] FIG. 2 depicts a diagram of a system for crowd level estimation in accordance with a first aspect of the present embodiment.

[0011] FIG. 3 depicts a diagram of a system for crowd level estimation in accordance with a second aspect of the present embodiment.

[0012] FIG. 4 depicts a flowchart of performance modelling in accordance with the second aspect of the present embodiment.

[0013] FIG. 5, comprising FIGs. 5A and 5B, depicts graphs of error distribution for crowd estimation in accordance with the second aspect of the present embodiment, wherein FIG. 5A depicts a graph of error distribution for crowd estimation of high

crowd level crowds and FIG. 5B depicts a graph of error distribution for crowd estimation of low crowd level crowds.

[0014] FIG. 6 depicts a graph of F-score variance for two crowd levels in accordance with the present embodiment.

[0015] FIG. 7 depicts a diagram of a system for crowd level estimation in accordance with a third aspect of the present embodiment.

[0016] FIG. 8 depicts a spatial pixel variation crowd estimation technique in accordance with the present embodiment.

[0017] FIG. 9 depicts automatic crowd estimation technique switching in accordance with the present embodiment.

[0018] FIG. 10 depicts a flowchart of a method for crowd level estimation in accordance with the present embodiment.

[0019] And FIG. 11 depicts a flowchart for selecting the best performing crowd estimation technique in accordance with the present embodiment.

[0020] Skilled artisans will appreciate that elements in the figures are illustrated for simplicity and clarity and have not necessarily been depicted to scale.

DETAILED DESCRIPTION

[0021] The following detailed description is merely exemplary in nature and is not intended to limit the invention or the application and uses of the invention. Furthermore, there is no intention to be bound by any theory presented in the preceding background of the invention or the following detailed description. It is the intent of the present embodiment to present methods and systems for real time robust

and optimized crowd estimation. When analyzed closely, it is possible to identify and/or model multiple techniques which can complement each other. In accordance with present embodiments, methods and systems to automatically switch between these crowd estimation techniques depending on a current crowd level (low crowd level, high crowd level) and other parameters tap these advantages to provide optimized crowd estimation. Note that throughout this description, high/low crowd density refers to a discrete crowd level and should not be considered as a measure of discrete crowd density.

[0022] Referring to FIG. 1A, an illustration 100 depicts a crowd 102 at a location 104 and a camera 106 arranged to capture images of the crowd 102 at the location 104. FIG. 1B depicts an illustration 150 of media 152 capturing images 154, 156. The images 154 are images of a high crowd level and the images 156 are images of a low crowd level.

[0023] Referring to FIG. 2, a diagram 200 depicts a system for crowd estimation in accordance with a first aspect of the present embodiment. The system includes an input module 202 for receiving an image of the crowd 102. In accordance with the present embodiments, a plurality of crowd estimation technique calculators 204 receive the image of the crowd 102 from the input module 202 and estimate crowd counts 205 therefrom. An equal plurality of performance modeling modules 206 are coupled to each of the crowd estimation technique calculators 204 for modeling each of the plurality of crowd estimation techniques based on an accuracy thereof at different crowd levels and/or at different locations.

[0024] A crowd estimation technique integration module 208 selects one or more of the plurality of crowd estimation techniques in response to the performance modeling thereof and an estimated crowd level and/or an estimated location. The crowd

estimation technique integration module 208 then estimates the crowd count of the crowd in the received image in accordance with the selected one or more of the plurality of crowd estimation techniques and outputs a final crowd count 210.

[0025] Accordingly, performance models of individual crowd estimation techniques are created at different crowd levels by using incoming image frames to generate the estimated crowd counts for the different crowd estimation techniques. Then, a crowd level estimation method determines which estimated crowd count to select or assign a high confidence value. In accordance with the present embodiment, the input module 202 can receive the image of the crowd and determine a region of interest within the image of the crowd. The crowd estimation technique integration module 208 then estimates one or both of the crowd level of the crowd in the region of interest within the image of the crowd or the location of the crowd in the region of interest within the image of the crowd.

[0026] Alternatively, the input module 202 can receive the image of the crowd and divide the image into a plurality of sub-regions. The input module could divide the image of the crowd into the plurality of sub-regions in consideration of a view point of the camera 106 which has captured the image or in respect of other criteria. The crowd estimation technique integration module 208 would then select one or more of the plurality of crowd estimation technique calculators 204 for each of the plurality of sub-regions in response to the performance modeling of the one or more of the plurality of crowd estimation techniques by the corresponding one of the plurality of performance modeling modules 206 and an estimated crowd level and/or an estimated location for the one of the plurality of sub-regions. The crowd estimation technique integration module 208 would then estimate the crowd count of the crowd in each of the plurality of sub-regions in accordance with the selected one or more of the

plurality of crowd estimation techniques for that one of the plurality of sub-regions. Finally, the crowd estimation technique integration module 208 would combine the estimated crowd counts for each of the plurality of sub-regions to obtain the final crowd count 210 of the crowd in the received image.

[0027] In accordance with the present embodiment, the plurality of performance modeling modules 206 could assign a real-time confidence value to each of the plurality of crowd estimation techniques in accordance with the performance modeling thereof. The system may then include a confidence value observer 212 coupled to the crowd estimation technique integration module 208 for removing one of the plurality of crowd estimation technique calculators 204 from selection when the real-time confidence value of the one of the plurality of crowd estimation techniques falls below a confidence value threshold.

[0028] The crowd estimation technique integration module 208 could further select multiple ones of the plurality of crowd estimation techniques 204 and combine the crowd estimation results 205 from the multiple crowd estimation techniques 204 to estimate the crowd count of the crowd in the received image. In accordance with the present embodiment, the crowd estimation technique integration module 208 can dynamically combine the crowd estimation results 205 from the multiple crowd estimation techniques 204 in accordance with the real-time confidence value thereof to estimate the final crowd count 210 of the crowd in the received image 102. The crowd estimation results 205 can be combined in accordance with an inverted weighted sum approach or in accordance with a normalized weighted sum approach.

[0029] A further enhancement of the system depicted in the diagram 200 could involve adding a foreground measurement module 214 coupled between the input module 202 and the crowd estimation technique integration module 208 to measure a

crowd level in a foreground of the image of the crowd to provide an estimated crowd level for use by the crowd estimation technique integration module 208 when selecting the one or more of the plurality of crowd estimation techniques 204.

[0030] Referring to FIG. 3, a diagram 300 depicts a system for crowd level estimation in accordance with a second aspect of the present embodiment. The system depicted in the diagram 300 implements performance modeling of crowd estimation techniques by the performance modeling modules 206 for each of the one or more crowd estimation techniques 204 determining a plurality of performances of the corresponding crowd estimation technique 204 at multiple crowd levels (e.g., HIGH crowd levels, LOW crowd levels) and modeling the performance of the crowd estimation technique 204 in response to the plurality of performances of the crowd estimation technique 204 at the multiple crowd levels.

[0031] This performance modeling operation of the performance modeling modules 206 is shown in a flowchart 400 of FIG. 4. Each of the performance modeling modules 206 collect 402 images at different crowd levels and categorize 404 those images into low crowd images and high crowd images. Each performance modeling module 206 then models 406 the performance of the corresponding crowd estimation technique 204 in response to the plurality of performances of the crowd estimation technique 204 at the different crowd levels.

[0032] Referring back to FIG. 3, the performance modeling modules 206 can also determine a plurality of performances of the corresponding crowd estimation technique 204 at locations of interest and model the performance of the crowd estimation technique 204 in response to the plurality of performances of the crowd estimation technique 204 at the locations of interest.

[0033] In accordance with the present embodiment, a performance modeling module 206 may model the performance of a corresponding crowd estimation technique 204 by determining an error distribution 302 of the plurality of performances of the crowd estimation technique, such as by determining an error of crowd counting for each of the plurality of performances and/or by determining a standard deviation of the error distribution for each of the plurality of performances of the crowd estimation technique, as an indicator of performance of the crowd estimation technique 204.

[0034] Referring to FIGs. 5A and 5B, graphs 500, 550 depict error distribution for crowd estimation in accordance with the second aspect of the present embodiment. The graph 500 depicts a graph of error distribution for crowd estimation of high crowd level crowds and the graph 550 depicts a graph of error distribution for crowd estimation of low crowd level crowds. Validation of an accuracy of the crowd estimation technique 204 with image samples at different crowd levels is used by the performance modeling module 206 in accordance with the present embodiment to generate the error distribution 302 at the considered crowd levels, the error referring to the deviation in the crowd estimation from the actual number of people.

[0035] The standard deviation (σ) of the error distribution 302 indicates the suitability of the crowd estimation technique 204. When the count estimate error is less, σ is small. For low crowd level σ as shown in the graph 550, the error distribution indicates the crowd estimation technique 204 has less error for low crowd levels as compared to high crowd levels (i.e., as shown in the distribution graph 500). The calculation of the standard deviation is shown in Equation (1) below.

$$\sigma = \sqrt{\frac{\sum_{i=1}^{M}(x_i)^2}{M-1}} \qquad (1)$$

where M is the number of samples, x_i is the error of the ith sample. The Equation (1) shows that if $\sigma_{low} \ll \sigma_{high}$, the particular crowd estimation technique 204 being performance modeled by the corresponding performance modeling module 206 is suitable for low crowd level estimation.

[0036] Referring back to FIG. 3, the performance modeling module 206 may alternatively model the performance of the corresponding crowd estimation technique 204 by determining an accuracy metric for the plurality of performances of the crowd estimation technique, wherein the accuracy metric may include a F-score and wherein the performance modeling module 206 determines the accuracy metric for the plurality of performances of the corresponding crowd estimation technique 204 by determining a variance of the F-score 304 for the plurality of performances of the crowd estimation technique. F-score is a measure of performance based on the number of humans not detected and other regions falsely detected as humans. The performance modeling module 206 may determine a F-score distribution 304 with respect to a mean of F-scores for the multiple performances of the crowd estimation technique 204 and then determine an indicator of performance of the crowd estimation technique 204 in response to both the variance of the F-score for the multiple performances of the crowd estimation technique 204 and the F-score distribution with respect to the mean of F-scores for the multiple performances of the crowd estimation technique 204.

[0037] FIG. 6 depicts a graph 600 showing F-score variance at a first crowd level 602 and a second crowd level 604. The F-score for samples at different crowd levels is used to find the F-Score variance at these crowd levels. The variance $V(Fcr)$ of F-scores at a particular crowd level can be calculated from Equation (2) below.

$$V(Fcr) = \frac{\sum_{i=1}^{M}\left(F_{cr}^{i} - \mu\right)^2}{M-1} \qquad (2)$$

where *cr* is low or high crowd level, M is the number of samples, μ is a mean of F-scores and F_{cr}^{i} is the F-score for the ith sample. If $V(F_{low}) \ll V(F_{high})$ and $\mu_{low} \gg \mu_{high}$, the particular crowd estimation technique 204 being performance modeled by the corresponding performance modeling module 206 is suitable for low crowd level estimation.

[0038] Referring to FIG. 7, a diagram 700 depicts a system for crowd level estimation in accordance with a third aspect of the present embodiment. In accordance with this third aspect, a crowd level estimation module 702 provides an estimated crowd level to the crowd estimation technique integration module 208 for use in selecting a most appropriate one of the crowd estimation techniques 204. The crowd level estimation module 702 can estimate a crowd level of the crowd in the input image received by the input module 202 in response to a crowd density level. This could be accomplished by focusing on a region of interest within the image. The input module 202 could receive the input image of the crowd and determine the region of interest within the input image of the crowd. Then, the crowd level estimation module 702 could estimate the crowd level of the crowd within the region of interest of the input image in response to the crowd density level at that region of interest.

[0039] The crowd level estimation module 702 may include a spatial pixel variation model building unit 704 for modeling spatial variations of each of a plurality of crowd levels in response to pixel density variations thereof to generate multiple models of crowd level spatial variations. The crowd level estimation module 702 can then estimate the crowd level for automatic crowd estimation technique switching 706 by determining a similarity of the input image of the crowd to each of the models of crowd level spatial variations built by the spatial pixel variation model building unit

704 and estimating the crowd level of the crowd in the input image in response to a most similar one of the models of crowd level spatial variations.

[0040] In regards to determining the most similar one of the models of crowd level spatial variations, the crowd level estimation module 702 can estimate the crowd level of the crowd in the input image in response to a probability density function of a similarity of the input image of the crowd and each of the plurality of models of crowd level spatial variations. More specifically, the crowd level estimation module estimates the crowd level of the crowd in the input image in response to a best fit model of the plurality of models of crowd level spatial variations as determined by the probability density function of the similarity of the input image of the crowd and each of the plurality of models of crowd level spatial variations.

[0041] The spatial pixel variation model building unit 704 can generate the plurality of models of crowd level spatial variations in response to one or more of a grayscale crowd histogram or a red-green-blue (RGB) crowd histogram 708, a crowd local binary pattern 710 or a crowd texture 712. The automatic crowd estimation technique switching 706 of the crowd level estimation module 702 can switch crowd estimation techniques in response to an estimated discrete level of the crowd in the input image.

[0042] Thus, crowd levels such as low crowd levels and high crowd levels are estimated to select or to assign higher confidence values to crowd estimation techniques 204 which perform better at the estimated crowd level. The crowd level estimation 702 is accomplished by first spatial pixel variation model building by the spatial pixel variation model building unit 704 and then automatic crowd estimation technique switching 706.

[0043] Referring to FIG. 8, a flow chart 800 and illustrations 820 depict the operation of the crowd level estimation module 702 where estimation is based on

modeling spatial variations of crowd levels by the spatial pixel variation model building unit 704. The flowchart 800 depicts the spatial variation modeling process in accordance with the present embodiment. At a location of interest 802, the camera 106 acquires 804 images of the crowd 102. Training images of the crowd are extracted 806 for required crowd levels (e.g., high crowd level or low crowd level). The spatial pixel variations are then extracted 808 from the training images and spatial pixel variation models are developed 810 for the required crowd levels.

[0044] Turning to the illustrations 820, each of the steps of the flowchart are shown pictorially. At an illustration 830, the camera 106 monitoring the location of interest 104 is selected. In an illustration 840, the video (media 152) is recorded from the location of interest 104 covering different crowd levels ranging from high crowd levels in images 154 to low crowd levels in images 156.

[0045] The illustrations 850 correspond to step 806 in the flowchart 800 where training images 852 for high crowd levels and training images 854 for low crowd levels are extracted. In this manner, image frames 852, 854 with different crowd levels covering a ‘no person case’ to ‘a fully crowded case’ are extracted from the video clip 152 recorded at step 804 in the flowchart 800.

[0046] At the next step 808, spatial pixel variations are extracted. A histogram approach for extracting spatial pixel variation is provided as an example in illustration 860. A grey scale histogram of an image is a frequency representation of the pixel intensities grouped at discrete pixel intensity levels called bins. Grey scale histograms 862, 864 of all the extracted images 852, 854 are recorded with 256 bins. The image-histogram pairs are grouped into high crowd level images 852 and histograms 862 and low crowd level images 854 and histograms 864 based on the number of humans in the images 852, 854.

[0047] At each crowd level, a bin-wise frequency averaging is performed considering all the image-histogram pairs. The averaging forms histogram models 872, 874 for each crowd level as pictured in the illustration 870. In operation, incoming image frames 154, 156 are compared against these histogram models 872, 874 to estimate a crowd level for each image frame.

[0048] Referring to FIG. 9, a flow chart 900 and illustrations 920 depict the operation of the crowd level estimation module 702 where estimation is based on automatic crowd estimation technique switching by the automatic crowd estimation technique switching 706 (FIG. 7). The flowchart 900 depicts the automatic crowd estimation technique switching process in accordance with the present embodiment. At a location of interest 902, the camera 106 acquires a live stream video 904 of images of the crowd 102. The spatial pixel variations are then extracted 906 from the acquired images and crowd level estimation is performed 908 by probability calculation based on similarity determination. Processing then selects or integrates 910 the appropriate crowd estimation technique 204.

[0049] Turning to the illustrations 920, each of the steps of the flowchart are shown pictorially. At an illustration 930, the camera 106 initiates a live video stream of the location of interest 104. In an illustration 940, the automatic periodic capture of image frames of the crowd 942 is initiated for the location of interest 104. The user can define an appropriate time interval for image frame capture 944.

[0050] At the next step 906, spatial pixel variations are extracted. An exemplary histogram approach for extracting spatial pixel variation 950 extracts a grey scale histogram 952 of an image as a frequency representation of the pixel intensities grouped at discrete pixel intensity levels.

[0051] At the next step 908, the histogram 952 is compared against all the histogram models 872, 874 generated in the model building stage 870 (FIG. 8). The histogram 952 is compared to the histogram model 872 at the illustration 962 and compared to the histogram model 874 at the illustration 964. The comparison is performed by calculating similarity scores between the histogram 952 of the incoming image frame and the histogram models 872, 874. Examples of the similarity calculation methods include the correlation method, the Bhattacharya distance method, the Chi-square method and the intersection method. The similarity calculation results in each similarity method acting as a classifier on whether the incoming image frame resembles a high crowd level 962 or a low crowd level 964.

[0052] For example, using four different pixel variation modeling methods with each method output model 872, 874 being compared against that of an incoming image frame 942 by four different similarity calculations would result in sixteen classifications. A Probability Density Function (PDF) can be constructed based on these sixteen classifications as shown in Equations (3) and (4).

$$P(X = high) = \frac{Total\ Number\ of\ High\ crowd\ level\ classifications}{16} \qquad (3)$$

$$P(X = low) = \frac{Total\ Number\ of\ low\ crowd\ level\ classifications}{16} \qquad (4)$$

[0053] At step 910, an incoming image frame 942 is classified to a particular crowd level based on the highest probability calculated at step 908. For crowd estimation technique selection at step 910, the crowd estimation technique 204 with the lowest σ or $V(Fcr)$ with a high Fcr at the estimated crowd level is selected.

[0054] For crowd estimation technique integration at step 910, the final count estimate (F_{count}) is calculated using Equation (5).

$$F_{count} = round\left(\frac{\sum_{i=1}^{N} Z_{ri} \times count_i}{N}\right) \qquad (5)$$

where i = 1 to N crowd estimation techniques and Z_{ri} is a re-weighted confidence value calculated by Equation (6).

$$Z_{ri} = 1 - Z_{normi} \quad (6)$$

where Z_{normi} is the normalized confidence value in the range [0,1] calculated using Equation (7).

$$Z_{normi} = \frac{Z_i - \min(Z)}{\max(Z) - \min(Z)} \quad (7)$$

where, Z can either be a set of σ or a set of $V(Fcr)$ generated for all the crowd estimation techniques 204. For example, where results from a first crowd estimation technique 204 (e.g., a combined head pattern estimation technique) and a second crowd estimation technique 204 (e.g., an individual head pattern estimation technique) can be represented as $Count_1$ and $Count_2$, Equation (8) shows mean averaging of the confidence values.

$$(Z_1 \text{ x } Count_1 + Z_2 \text{ x } Count_2)/2 \quad (8)$$

where $Z_1 >> Z_2$ and $Z_1 + Z_2 = 1$ (for example Z_1 could be 1 and Z_2 could be 0).

[0055] Referring to FIG. 10, a flowchart 1000 depicts a method for crowd level estimation in accordance with the present embodiment. When an input image of a crowd is received 1002, each crowd estimation technique is applied 1004 to the image and crowd counts are calculated 1006 for each of the crowd estimation techniques. At the same time, spatial pixel variations are extracted 1008 from the received input image. The extracted spatial pixel variations are compared 1010 against spatial pixel models to find the highest similarity as described above. The crowd level is determined from the comparisons 1010 and confidence values are assigned 1012. The counts from the crowd estimation techniques 1006 are then integrated 1014 with the crowd level/confidence values 1012 to estimate the final crowd count 1016.

[0056] Methods in accordance with the present embodiment can also be used to select the best performing crowd estimation technique. In this case, the incoming image frame is not processed by all crowd estimation techniques; only the selected techniques will process the incoming image frame. Referring to FIG. 11, a flowchart 1100 depicts this selection process. The input image is received 1102 and spatial pixel variations are extracted 1104. The spatial pixel variations are compared 1106 against all spatial pixel models to determine the crowd level and select the crowd estimation technique for that crowd level 1108. The selected crowd estimation technique is applied 1110 to estimate the final crowd count 1112.

[0057] Thus, it can be seen that the present embodiment provides methods and systems for real time robust and optimized crowd estimation. When analyzed closely, it is possible to identify and/or model multiple techniques which can complement each other. In accordance with present embodiments, methods and systems to automatically switch between these crowd estimation techniques depending on a current crowd level (low crowd level, high crowd level) and other parameters tap these advantages to provide optimized real time crowd estimation with improved accuracy in a variety of crowd conditions and crowd locations.

[0058] While exemplary embodiments have been presented in the foregoing detailed description of the invention, it should be appreciated that a vast number of variations exist. It should further be appreciated that the exemplary embodiments are only examples, and are not intended to limit the scope, applicability, operation, or configuration of the invention in any way. Rather, the foregoing detailed description will provide those skilled in the art with a convenient road map for implementing an exemplary embodiment of the invention, it being understood that various changes may be made in the function and arrangement of steps and method of operation

described in the exemplary embodiment without departing from the scope of the invention as set forth in the appended claims.

CLAIMS

What is claimed is:

1. A method for crowd level estimation comprising:

receiving an input image of a crowd; and

estimating a crowd level of the crowd in the input image in response to a crowd density level.

2. The method in accordance with Claim 1 wherein estimating the crowd level comprises:

determining a region of interest within the input image of the crowd; and

estimating a crowd level of the crowd within the region of interest of the input image in response to a crowd density level.

3. The method in accordance with either Claim 1 or Claim 2 wherein estimating the crowd level in response to the crowd density level comprises:

determining a similarity of the input image of the crowd to each of a plurality of models of crowd level spatial variations; and

estimating the crowd level of the crowd in the input image in response to a most similar one of the plurality of models of crowd level spatial variations.

4. The method in accordance with Claim 3 wherein determining the similarity of the input image of the crowd to each of the plurality of models of crowd level spatial variations comprises modeling spatial variations of each of a plurality of

crowd levels in response to pixel density variations thereof to generate the plurality of models of crowd level spatial variations.

5. The method in accordance with either Claim 3 or Claim 4 wherein estimating the crowd level of the crowd in the input image comprises estimating the crowd level of the crowd in the input image in response to a probability density function of a similarity of the input image of the crowd and each of the plurality of models of crowd level spatial variations.

6. The method in accordance with Claim 5 wherein estimating the crowd level of the crowd in the input image further comprises estimating the crowd level of the crowd in the input image in response to a best fit model of the plurality of models of crowd level spatial variations as determined by the probability density function of the similarity of the input image of the crowd and each of the plurality of models of crowd level spatial variations.

7. The method in accordance with either Claim 5 or Claim 6 wherein the plurality of models of crowd level spatial variations comprises a model generated in response to a grayscale crowd histogram.

8. The method in accordance with either Claim 5 or Claim 6 wherein the plurality of models of crowd level spatial variations comprises a model generated in response to a red-green-blue (RGB) crowd histogram.

9. The method in accordance with either Claim 5 or Claim 6 wherein the plurality of models of crowd level spatial variations comprises a model generated in response to a crowd local binary pattern.

10. The method in accordance with either Claim 5 or Claim 6 wherein the plurality of models of crowd level spatial variations comprises a model generated in response to crowd texture.

11. The method in accordance with any of the proceeding claims wherein estimating the crowd level comprises estimating a discrete level of the crowd in the input image.

12. The method in accordance with any of the proceeding claims wherein receiving the image of the crowd comprises:

receiving the image of the crowd; and

dividing the image of the crowd into a plurality of sub-regions, and

wherein estimating the crowd level of the crowd in the received image comprises estimating the crowd level of the crowd in each of the plurality of sub-regions in response to the crowd density for that one of the plurality of sub-regions.

13. A system for crowd level estimation comprising:

an input module for receiving an input image of a crowd; and

a crowd level estimation module for estimating a crowd level of the crowd in the input image in response to a crowd density level.

14. The system in accordance with Claim 13 wherein the input module receives the input image of the crowd and determines a region of interest within the input image of the crowd, and

wherein the crowd level estimation module estimates the crowd level of the crowd within the region of interest of the input image in response to the crowd density level.

15. The system in accordance with either Claim 13 or Claim 14 wherein the crowd level estimation module includes a spatial pixel variation model building unit for modeling spatial variations of each of a plurality of crowd levels in response to pixel density variations thereof to generate a plurality of models of crowd level spatial variations, and wherein the crowd level estimation module estimates the crowd level by determining a similarity of the input image of the crowd to each of the plurality of models of crowd level spatial variations and estimating the crowd level of the crowd in the input image in response to a most similar one of the plurality of models of crowd level spatial variations.

16. The system in accordance with Claim 15 wherein the crowd level estimation module estimates the crowd level of the crowd in the input image in response to a probability density function of a similarity of the input image of the crowd and each of the plurality of models of crowd level spatial variations.

17. The system in accordance with Claim 16 wherein the crowd level estimation module estimates the crowd level of the crowd in the input image in response to a best fit model of the plurality of models of crowd level spatial variations

as determined by the probability density function of the similarity of the input image of the crowd and each of the plurality of models of crowd level spatial variations.

18. The system in accordance with any of Claims 15 to 17 wherein the spatial pixel variation model building unit generates the plurality of models of crowd level spatial variations in response to one or more of a grayscale crowd histogram, a red-green-blue (RGB) crowd histogram, a crowd local binary pattern or a crowd texture.

19. The system in accordance with any of Claims 13 to 18 wherein the crowd level estimation module estimates a discrete level of the crowd in the input image.

20. The system in accordance with any of Claims 13 to 19 wherein the input module receives the image of the crowd and divides the image of the crowd into a plurality of sub-regions, and

wherein the crowd level estimation module estimates the crowd level of the crowd in each of the plurality of sub-regions in response to the crowd density for that one of the plurality of sub-regions.

21. A computer readable medium having instructions stored thereon which when compiled onto a processor perform the method of:

receiving an input image of a crowd;

determining a region of interest within the input image of the crowd;

determining a similarity of the region of interest within the input image of the crowd to each of a plurality of models of crowd level spatial variations, each of the plurality of models of crowd level spatial variations generated by modeling spatial variations of each of a plurality of crowd levels in response to pixel density variations thereof; and

estimating the crowd level of the crowd in the region of interest within the input image in response to a most similar one of the plurality of models of crowd level spatial variations.

22. The computer readable medium in accordance with Claim 19 wherein the instructions stored thereon, when compiled onto a processor, will generate at least some of the plurality of models of crowd level spatial variations by modeling spatial variations of each of the plurality of crowd levels in response to one or more of a grayscale crowd histogram, a red-green-blue (RGB) crowd histogram, a crowd local binary pattern or a crowd texture.