Optical Engineering 46(4), 047201 (April 2007)

Disjoint track matching based on a major color spectrum histogram representation

Eric Dahai Cheng
Massimo Piccardi
University of Technology, Sydney (UTS)
Faculty of Information Technology
P.O. Box 123, Broadway
NSW 2007, Australia
E-mail: Eric.Cheng@eng.uts.edu.au

Abstract. A disjoint track matching algorithm is proposed based on major color spectrum histogram representation (MCSHR) matching and post-matching integration algorithms; this algorithm is useful for reconnecting broken tracks due to occlusions and potentially useful for tracking a single object across multiple, disjoint cameras. An incremental MCSHR (IMCSHR) matching algorithm is also proposed to cope with small pose changes occurring along the track. First, an MCSHR is introduced to represent a moving object by its most frequent colors. Then a two-directional (2-D) similarity measurement algorithm based on the most similar major color searching algorithm is proposed to measure the similarity of the two given moving objects by using IMCSHR in a few frames. Last, our track matching algorithm extends the multiframe matching along the object's tracks by a post-matching integration algorithm. Experimental results have shown that the similarity of two tracks from the same moving objects has proven as high as 89 to 97%, while the similarity of two tracks from different moving objects has been kept as low as 14 to 54%. The post-matching integration algorithm has been proven to be able to make track matching more robust and reliable.
 [DOI: 10.1117/1.2730482]

Subject terms: moving object tracking matching; major color spectrum histogram representation (MCSHR); color distance; similarity measurement.

Paper 060678 received Sep. 4, 2006; accepted for publication Oct. 2, 2006; published online Apr. 18, 2007.

1 Introduction

Video surveillance plays an important role in anti-terrorism efforts, especially after September 11, 2001. Moving object tracking through a network of cameras is an important function for effective video surveillance of wide areas, and matching of moving objects in disjoint cameras is becoming more and more important with the increased use of networks of cameras in surveillance systems.[1–17] In traditional tracking methods, people assume the continuity in space and time in successive observation frames, but in most cases, the observations of the same object are separate in time and space, and the requirement of camera calibration or complete site modeling is not available. The tracking of a moving object in disjoint tracks and across multiple and disjoint cameras is becoming an urgent research area.[18,19] On the other hand, matching of moving objects in disjoint tracks and disjoint cameras is a challenging task, since no continuous information is provided in this case. The assumption in this work is that tracks are available from within single camera views, and the goal is to find correspondences between such tracks.

Accordingly, in this paper we present a track matching algorithm based on incremental major color spectrum histograms[20–22] matching and post-matching integration. First, a color distance based on a normalized geometric distance between two points in the Red, Green, Blue (RGB) space is introduced to measure the similarity of any two colors. Then, by using the color distance and a given threshold, all pixels from a moving object MOi in a given frame t are clustered into a limited number of major colors, with each color's frequency defined as the number of pixels with that color. Such colors are then sorted in descending frequency order, and the first certain percentage (for example, 90%) of colors is used to represent the moving object. We call this histogram the major color spectrum histogram representation (MCSHR) of MOi, t. Given two arbitrary moving objects MOi, t and MOj, u from two different frames t and u, a similarity criterion based on the

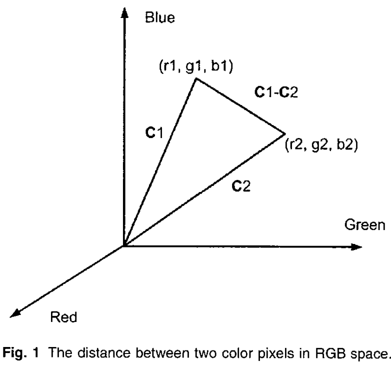


Fig. 1 The distance between two color pixels in RGB space.

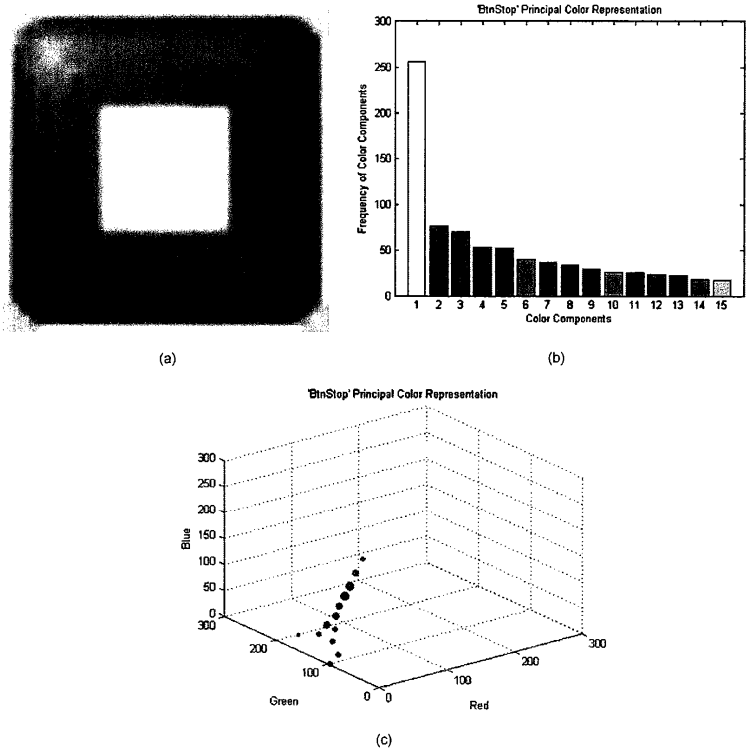


Fig. 2 Major color histogram representation of the BtnStop sign: (a) BtnStop sign picture, (b) major color representation histogram of BtnStop with the color distance threshold of 0.01, and (c) major color representation histogram of BtnStop in RGB Space. (Color online only.)

major color representation is proposed to assess their matching (single-frame matching).[23,24] In order to deal with pose changes along the track, an incremental major color spectrum histogram algorithm is proposed here in this paper, in which the similarity is calculated based on the incremental MCSHR (IMCSHR), which is close to the similarity of single-frame matching. The IMCSHR matching is then extended along the two moving objects' tracks by selecting the same number of frames in each track, performing the matching between the corresponding frames of each track, and integrating the matching results over time. Last, the time-integrated decision is compared against an assigned threshold to provide the final track matching decision. To the best of our knowledge, this is one of the first papers in the current literature to tackle the problem of track matching across disjoint tracks or disjoint camera views.[18,19] Different from those previous papers, our approach does not require global track matching[18] or rely on a topographic model of the camera network.[19] Experimental results proved that with a few (typically, three) frames of MCSHR integration, the proposed IMCSHR algorithm can make matching more robust and reliable than single-frame matching, especially for small pose changes, and the similarity measurement calculation can be saved up to 60% by cutting 10% of the major colors at its tail of major color representation, since these lower-frequency colors take more color bins.

2 Major Color Spectrum Histogram

2.1 *Concept of Color Distance*

If we use 1 byte to represent each color, there are 16.8 million (16,777,216) colors in the total RGB space. It is very difficult to compare two objects based on so many possible colors. In this paper, we introduce a concept of "color distance" between two color pixels, which is based

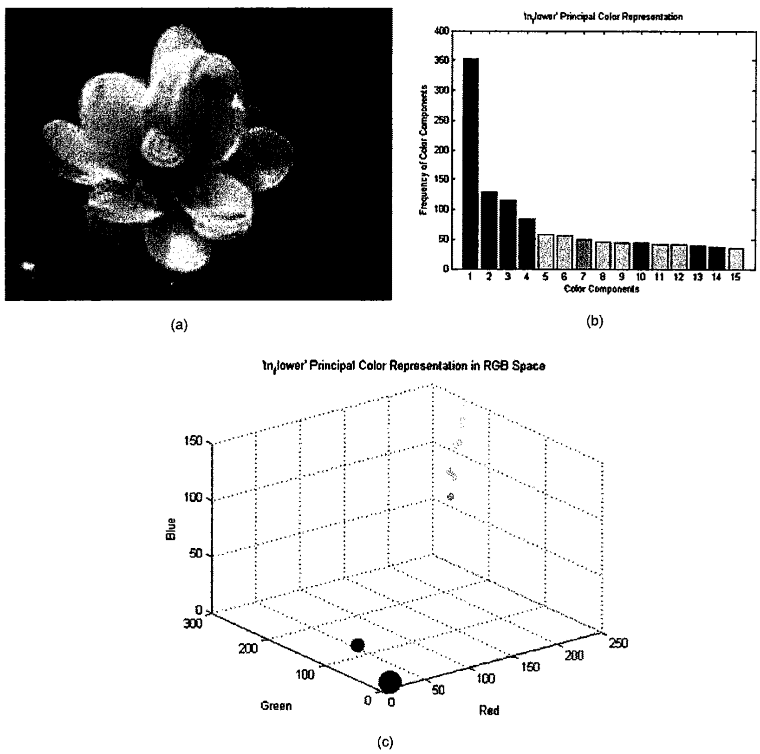


Fig. 3 Major color histogram representation of the tn_flower picture: (a) tn_flower picture, (b) major color representation histogram of tn_flower with the color distance threshold of 0.01, and (c) major color representation histogram of tn_flower in RGB space. (Color online only.)

on the normalized geometric distance between two color pixels in RGB space. The shorter the color distance between two colors, the more similar they are. Such a geometric distance is defined in Eq. (1) and exemplified in Fig. 1:

$$d(C_1,C_2)=\frac{\|C_1-C_2\|}{\|C_1\|+\|C_2\|}=\frac{[(r_1-r_2)^2+(g_1-g_2)^2+(b_1-b_2)^2]^{1/2}}{(r_1^2+g_1^2+b_1^2)^{1/2}+(r_2^2+g_2^2+b_2^2)^{1/2}}, \qquad (1)$$

where C_1 and C_2 are color vectors, as shown in Fig. 1.

2.2 *Moving Object MCSHR*

By using the concept of color distance, we can scale down the possible colors from 16.7 million to a very limited number of "major colors" (15 to 300, depending on the size and distribution of colors in the moving object) without losing much accuracy in representing the moving object. This theory is true because, in most cases, there are always several major colors in a moving object, so the colors that rarely appear in the moving object are discarded, and those colors that are similar are merged into one major color.[17–19] Examples of MCSHR under the color scaling threshold of 0.01 are shown in Figs. 2 and 3.

A BtnStop sign picture is shown in Fig. 2(a). In this picture, we can see that the colors that most often appear are white (or close to white) and green (or close to green). Figure 2(b) shows a histogram of the major colors of the BtnStop sign, in which we can see that there are more than 250 white (or close to white—the distance is less than 0.01) colors in color bar 1. The remaining green colors are distributed in color bars 2 to 15, with approximate pixel numbers of 75 to 10. Figure 2(c) shows the positions of these major colors in the RGB space, in which the sizes of the color circles represent the frequency of each color. We can

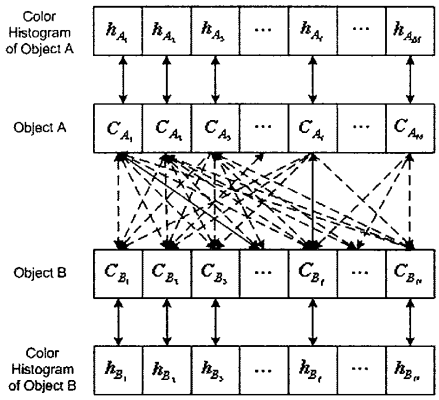


Fig. 4 The most-similar-color searching algorithm.

see that the green circles are close to the Green axis. The white circles are far from the Red, Blue, and Green axes and cannot be seen here.

Another example of a major color histogram representation is shown in Fig. 3. A tn_flower picture is shown in Fig. 3(a). In this picture, we can see that the colors that appear most often are dark black (or close to black) and yellow (or close to yellow). Figure 3(b) shows a histogram of the major colors of the tn_flower picture, in which we can see that there are four black dark color spectrum bins under the color distance threshold of 0.01. The numbers of black or dark pixels that are dropped in these bins are about 350, 125, 120, and 85, respectively (or close to dark—any pixels whose color distance from the bin reference color is less than 0.01). The yellow colors are distributed in color spectrum bins 5, 6, 7, 8, 9, 11, 12, and 15 with color changes more than the color distance threshold 0.01. The numbers of pixels of the yellow colors are between about 60 and 40. There are also three black or dark colors spread in bins 10, 13, and 14, with pixel numbers between 40 and 35. Figure 3(c) shows the positions of these major colors in the RGB space, in which the size of these color circles represent the frequency of each color. We can see that the dark circles are close to the RGB axes' zero-center, and the yellow circles are far from the RGB axes' zero-center.

3 Matching Algorithms

3.1 *Moving Objects Similarity Measurements*

In this section, a similarity measurement based on a most-similar-color searching algorithm is proposed to measure the similarity between two moving objects. This algorithm is based on the major color spectrum histogram of the two moving objects.[17–19] We assume that there are M colors in the major color histogram spectrum of moving object A, which can be represented as:

$$MCSHR(A)=\{C_{A_1},C_{A_2},\ldots,C_{A_M}\}, \tag{2}$$

where C_{A_i}, $i=1,2,\ldots,M$ is the color vector (RGB) of the major colors in object A. Object A's corresponding major color frequencies can be represented as:

$$p(A)=\{p(A_1),p(A_2),\ldots,p(A_M)\}. \tag{3}$$

Similarly, the major color spectrum histogram of object B can be represented as follows:

$$MCSHR(B)=\{C_{B_1},C_{B_2},\ldots,C_{B_N}\}, \tag{4}$$

$$p(B)=\{p(B_1),p(B_2),\ldots,p(B_N)\}. \tag{5}$$

In order to define the similarity between two moving objects, a subset of $MCSHR(B)$ in Eq. (4) is first defined as:

$$MCSHR'(B)=\{C_{B'_1},C_{B'_2},\ldots,C_{B'_L}\}, \tag{6}$$

where the distance between $C_{B'_j}$, $j=1,2,\ldots,L$ and C_{A_i} is less than the given threshold σ. This means that all colors in Eq. (6) are similar to color C_{A_i}.

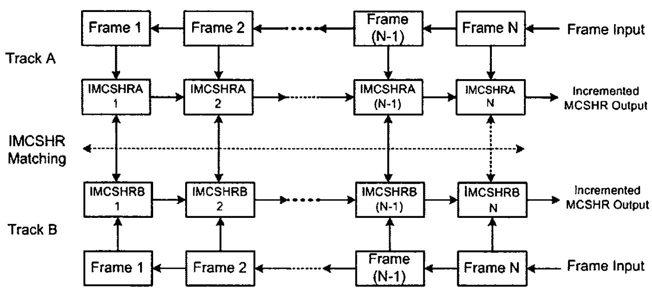


Fig. 5 Multiframe incremental major color spectrum histogram representation and matching algorithms.

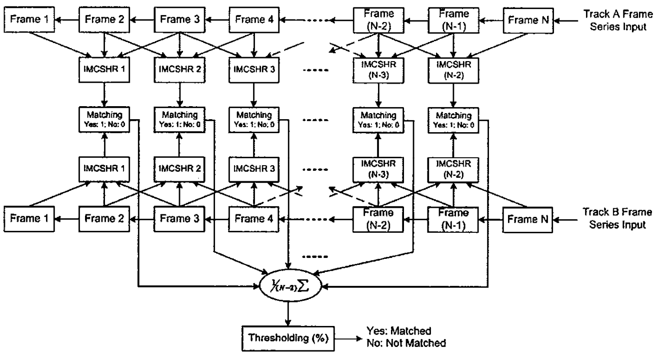


Fig. 6 Multi-frame, incremental MCSHR matching and post-matching integration algorithm.

$C_{B_j|A_i}$ is then defined as the most-similar color component of C_{A_i} in object B, i.e., the member of subset $MCSHR'(B)$ in Eq. (6) that satisfies Eq. (7):

$$d(C_{B_j|A_i}, C_{A_i}) = \min\{d(C_{B'_1}, C_{A_i}), d(C_{B'_2}, C_{A_i}), \ldots, d(C_{B'_L}, C_{A_i})\}. \tag{7}$$

The most-similar-color searching algorithm is shown in Fig. 4.

The portion of C_{A_i} in object A can be calculated with the following equation:

$$p_{norm}(A_i) = \frac{p(A_i)}{\sum_{i=1,2,\ldots,M} p(A_i)}. \tag{8}$$

Similarly, the portion of the corresponding color of C_{A_i} in object B can be calculated with the following equation:

$$p_{norm}(B_j|A_i) = \frac{p(B_j|A_i)}{\sum_{j=1,2,\ldots,N} p(B_j)}. \tag{9}$$

The similarity of color C_{A_i} in object A with its corresponding color C_{B_j} in object B is then defined as:

$$Similarity(C_{A_i}, C_{B_j}) = \min\left\{\frac{p(A_i)}{\sum_{i=1,2,\ldots,M} p(A_i)}, \frac{p(B_j|A_i)}{\sum_{j=1,2,\ldots,N} p(B_j)}\right\}. \tag{10}$$

The similarity of object A and object B based on the most-similar-color searching algorithm in the direction from A to B is defined as:

$$\begin{aligned} Similarity(A,B) &= \sum_{i=1}^{M} Similarity(C_{A_i}, C_{B_j}) \\ &= \sum_{i=1}^{M} \min\left\{\frac{p(A_i)}{\sum_{i=1,2,\ldots,M} p(A_i)}, \frac{p(B_j|A_i)}{\sum_{j=1,2,\ldots,N} p(B_j)}\right\}. \end{aligned} \tag{11}$$

Similarly, the similarity of color C_{B_j} in object B with its corresponding color C_{A_i} in object A is defined as:

$$Similarity(C_{B_j}, C_{A_i}) = \min\left\{\frac{p(B_j)}{\sum_{j=1,2,\ldots,N} p(B_j)}, \frac{p(A_i|B_j)}{\sum_{i=1,2,\ldots,M} p(A_i)}\right\}. \tag{12}$$

Thus, the similarity of object B and object A based on the most-similar-color searching algorithm in the direction from B to A is defined as:

$$\begin{aligned} Similarity(B,A) &= \sum_{j=1}^{N} Similarity(C_{B_j}, C_{A_i}) \\ &= \sum_{j=1}^{N} \min\left\{\frac{p(B_j)}{\sum_{j=1,2,\ldots,N} p(B_j)}, \frac{p(A_i|B_j)}{\sum_{i=1,2,\ldots,M} p(A_i)}\right\}. \end{aligned} \tag{13}$$

In order to derive the similarity of objects A and B, the minimum and maximum of Eqs. (11) and (13) are defined in Eqs. (14) and (15), respectively, as:

$$Similarity_{min} = \min\{Similarity(A,B), Similarity(B,A)\}, \tag{14}$$

$$Similarity_{max} = \max\ \{Similarity(A,B), Similarity(B,A)\}. \tag{15}$$

Eventually, we combine such minimum and maximum similarities into a single final value, $Similarity_{A,B}$. If $Similarity_{min}$ is less than a given discrimination threshold $\eta_{discrim}$, the similarity of object A and object B is simply defined as:

$$Similarity_{A,B} = \min\ \{Similarity(A,B), Similarity(B,A)\}. \tag{16}$$

The rationale is that in this case the two similarities between A and B, Eqs. (11) and (13), are very asymmetric, and for this reason, we decide to bound them by the lowest value. If $Similarity_{min}$ is instead above or equal to the discrimination threshold $-\eta_{discrim}$, we define:

$$Similarity_{A,B} = 1 - \frac{Similarity_{\max} - Similarity_{\min}}{Similarity_{\max} + Similarity_{\min}}. \tag{17}$$

In this case, we are confident that the two visual objects are possibly the same physical one. As a further verification, we choose to check the difference between the maximum and minimum similarities in ratio form, as shown in Eq. (17), in which we can see that the bigger the difference between maximum and minimum similarity, the less similar are the two tracks. Eventually, matching is assessed if $Similarity_{A,B}$ is above an assigned similarity threshold.

3.2 *Multiframe IMCSHR*

In order to cope with small pose changes occurring along the track, a multiframe IMCSHR and track matching algorithms are proposed here, in which the same number of frames from both tracks are chosen, as shown in Fig. 5.

With the incremental major color representation, the major colors of frame F_i are computed not only on the frame itself, but also on the window of the last N frames $\{F_{i-N}, F_{i-N+1}, \ldots, F_{i-1}, F_i\}$. First, the major colors of frame F_{i-N} are computed as described in Section 2. Then, instead of computing the major colors of frame F_{i-N+1} again from scratch (i.e., starting from an empty table), we compute them starting from those of frame F_{i-N}. We proceed with the other frames in a similar way to eventually obtain the IMCHSR of frame F_i. For those frames with index $j<N$ in the initial part of the track, we restrict the incremental computation to the available frames only $\{F_1, \ldots, F_j\}$. (Alternatively, such frames could be simply dismissed, but we prefer a partial computation so as to prevent dismissing a significant part of the track.) So, in Fig. 5, the matching between IMCSHR 1 in tracks A and B is actually the single-frame MCSHR matching because there is no IMCSHR available yet. The matching between IMCSHR N in tracks A and B is instead an N-frame integrated MCSHR matching. In order to study the effects of the integrated MCSHR on the matching process, the matchings were carried out at each stage of MCSHR integration, as shown in Fig. 5.

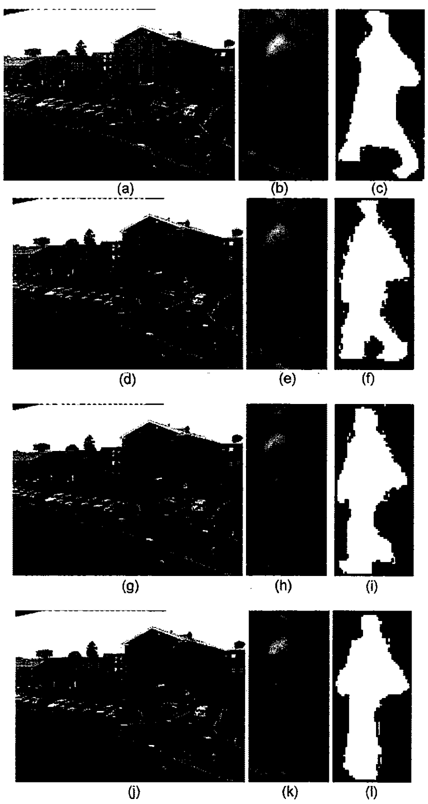


Fig. 7 Moving object examples for incremental MCSHR test: (a) frame 2040, (b) moving object, (c) mask, (d) frame 2042, (e) moving object, (f) mask, (g) frame 2044, (h) moving object, (i) mask, (j) frame 2046, (k) moving object, and (l) mask.

3.3 *Multiframe Incremental Major Color Spectrum Histogram Matching and Post-Matching Integration Algorithm*

A multiframe incremental major color spectrum histogram matching and post-matching integration algorithm is proposed in this section, as shown in Fig. 6. The matching is based on an IMCSHR, which makes the object representation more robust with respect to small pose changes. The post-matching integration over multiple frames makes the final track matching more robust and confident.

In the following, we assume that objects A and B are tracked over a sequence of N frames each (tracks A and B, respectively). The major colors of track A at the i'th frame are represented as:

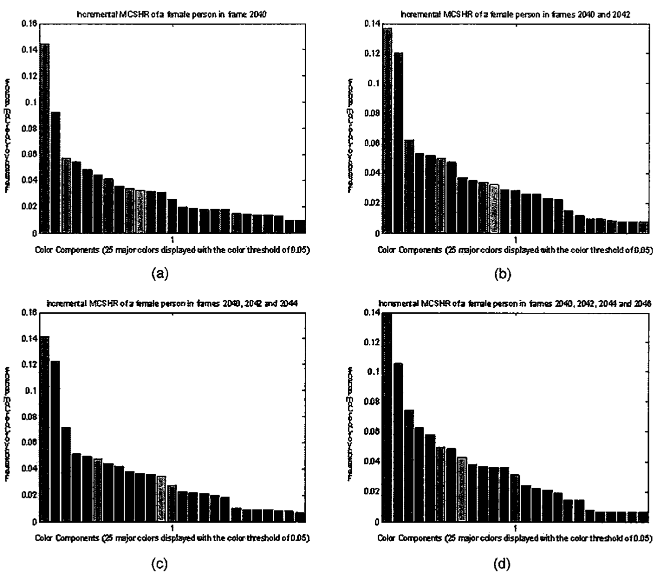


Fig. 8 Examples of incremental MCSHR: (a) incremental MCSHR of frame 2040; (b) incremental MCSHR of frames 2040 and 2042; (c) incremental MCSHR of frames 2040, 2042, and 2044; and (d) incremental MCSHR of frames 2040, 2042, 2044, and 2046. (Color online only.)

$$MCSHR(A_i) = \{C_1(A_i), C_2(A_i), \ldots, C_{M_i}(A_i)\}, \tag{18}$$

with corresponding frequencies:

$$p(A_i) = \{p_1(A_i), p_2(A_i), \ldots, p_{M_i}(A_i)\}. \tag{19}$$

The *incremental* major color spectrum histogram of the i'th frame in track A can then be represented as:

$$IMCSHR(A_i) = \sum_{j=i}^{i+K} MCSHR(A_j), \tag{20}$$

$$p_{IMCSHR}(A_i) = \sum_{j=i}^{i+k} p(A_j). \tag{21}$$

The Σ sign in Eqs. (20) and (21) is used to mean a special "summation," i.e., the merging accumulation of the MCSHRs of frames $i, \ldots, i+k$ based on the color threshold. The k is typically chosen between 2 and 4 (corresponding to three to five frames of integration).

Similarly, the major color spectrum in track B in the i'th frame is given by:

$$MCSHR(B_i) = \{C_1(B_i), C_2(B_i), \ldots, C_{N_i}(B_i)\}, \tag{22}$$

$$p(B_i) = \{p_1(B_i), p_2(B_i), \ldots, p_{N_i}(B_i)\}. \tag{23}$$

Accordingly, the incremental major color spectrum histogram in the i'th frame in track B is represented as:

$$IMCSHR(B_i) = \sum_{j=i}^{i+K} MCSHR(B_j), \tag{24}$$

$$p_{IMCSHR}(B_i) = \sum_{j=i}^{i+K} p(B_j). \tag{25}$$

Similarity between the object of track A at frame i and that of track B at frame i, $Similarity[int(A_i), int(B_i)]$, is then given by the same similarity measure in Eqs. (16) and (17) by simply replacing Eqs. (2) to (5) with Eqs. (20), (21), (24), and (25). Matching between the two frames, $Matching[int(A_i), int(B_j)]$, is stated if the similarity crossed the assigned similarity threshold.

In order to assess the matching between the two tracks A and B as a whole, we compare pairs of frames from the two tracks in track order, make a decision on their matching, and finally integrate the decision along N frames. Post-matching integration is used to make the matching decision more robust; it can be represented as:

Table 1 Incremental MCSHR matching with respect to the number of major colors.

No. of IMCSHR	3	
Frame index no.	2048–2052	2158–2162
Track no.	1	2
Similarity (MC)	0.9266	
Similarity (MC)	0.9663	
Similarity (MC)	0.9742	

$$PostMatchingInt = \sum_{i=1}^{N-K} Matching[\mathrm{int}(A_i), \mathrm{int}(B_i)]. \quad (26)$$

Based on Eq. (26), if *PostMatchingInt* is greater than a final percentage threshold η_{final}, we say that the two tracks are matched; otherwise, they are unmatched.

4 Experimental Results and Analysis

There are three parts in the experiment section, in which we use presegmented images, which include extracted moving objects and extracted moving object masks that have been obtained manually to avoid segmentation errors. There are five tracks of data that have been used in our experiments, in which five moving objects have been detected and extracted, including two female persons, a male person, a blue car, and a white van. In Sec. 4.1, we present a multiframe IMCSHR and performance test against the number of incremental frames and the number of major colors. In Sec. 4.2, we present matching tests on the same moving objects, which include a female person in frames 2040 to 2052 versus the same person in frames 2150 to 2162, and a blue car in frames 0510 to 0522 versus the same car in frames 0600 to 0612. In Sec. 4.3, we present matching tests on different moving objects, which include a female person in frames 0400 to 0412 versus another female person in frames 2150 to 2162, a male person in frames 0850 to 0862 versus a female person in frames 2150 to 2162, and a blue car in frames 0510 to 0522 versus a white van in frames 0700 to 0712.

Table 2 Incremental MCSHR matching with respect to the number of incremental frames.

Test case	Frame no.	Track no.	No. of frames	Similarity	Matching results
1	2044	1	1	0.9965	1 (Yes)
	2154	2			
2	2044–2046	1	2	0.9657	1 (Yes)
	2154–2156	2			
3	2044–2048	1	3	0.9650	1 (Yes)
	2154–2158	2			
4	2044–2050	1	4	0.9426	1 (Yes)
	2154–2160	2			
5	2044–2052	1	5	0.9422	1 (Yes)
	2154–2162	2			

4.1 *Multiframe IMCSHR and Performance Tests*

4.1.1 *Multiframe IMCSHR*

The proposed multiframe IMCSHR is calculated and tested here on the data in which a moving female person has been detected and tracked in frames 2040, 2042, 2044, and 2046. The moving object in these frames and their major color spectrum histograms (IMCSHR) are shown in Figs. 7 and 8, respectively.

Figure 8 shows that with the increase in the number of integration frames (from 2 to 4), the MCSHR tends to become more stable, in the sense that the actual major colors of the person (only 25 major color bins in the histograms shown in Fig. 8) are emphasized with respect to the other colors. Practically, because there are always errors in the process of moving object detection, together with small pose changes, an integration over a few frames (around 3 or 4) proved helpful for a more accurate and robust moving object representation. Further increase in the number of integrated frames does not necessarily improve the representation accuracy, as shown in this section and later experiments.

Table 3 Results of IMCSHR matching and post-integration.

Test case	Frame no.	Track no.	Similarity	Matching results
1	2040–2044	1	0.8899	1 (Yes)
	2150–2154	2		
2	2042–2046	1	0.9404	1 (Yes)
	2152–2156	2		
3	2044–2048	1	0.9650	1 (Yes)
	2154–2158	2		
4	2046–2050	1	0.9695	1 (Yes)
	2156–2160	2		
5	2048–2052	1	0.9742	1 (Yes)
	2158–2162	2		
Post-integration	2040–2052	1	N/A	100% (Yes)
	2150–2162	2		

Note: With 90% of major colors cut off, color threshold=0.05, similarity color threshold=0.05, discrimination threshold=0.55, IMCSHR matching threshold=0.8, and final integration matching threshold =0.8.

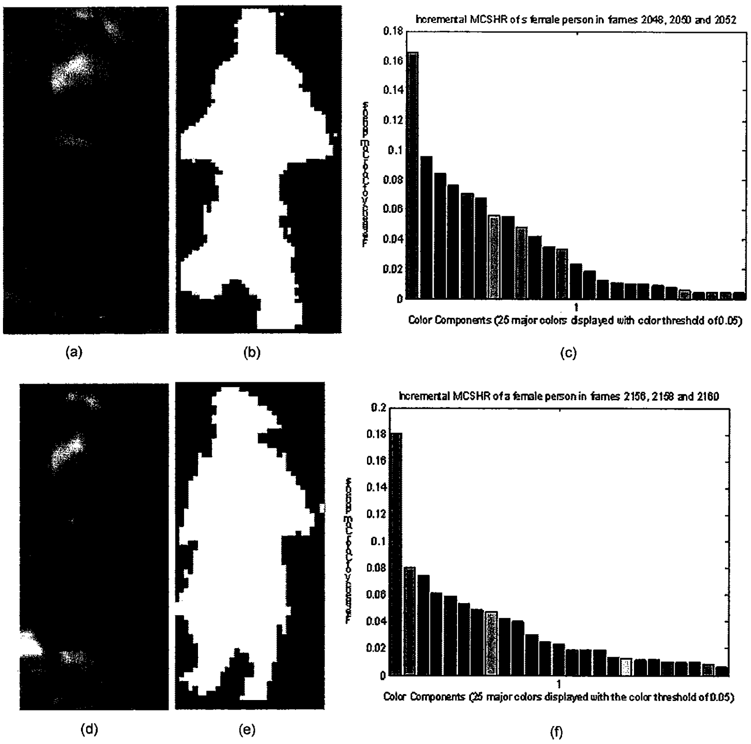


Fig. 9 Moving objects (MO) from track 1, frames 2040 to 2052, and track 2, frames 2150 to 2162: (a) MO in frame; (b) mask; (c) IMCSHR of 2048, 2050, and 2052; (d) MO in frame; (e) mask; and (f) IMCSHR of 2156, 2158, and 2160. (Color online only.)

4.1.2 *Matching performance versus the number of major colors*

As stated earlier, the number of major colors used in the representation can have a strong influence on the matching performance. Table 1 gives evidence for this by showing the matching results of a three-frame incremental MCSHR with respect to the change in the number of major colors for the color threshold and the similarity threshold of 0.05. The results in Table 1 show that the matching rate has increased from 93% to 97% with the increase of the number of major colors from 70% to 90%, while the amount of matching calculation is increased with the increase of the number of bins. Based on our test, 90% IMCSHR is considered a good trade-off between accuracy and computational load.

4.1.3 *Matching performance versus the number of incremental frames*

The number of incremental fames used in the representation can influence the matching performance. Table 2 gives evidence to this by showing the matching results of a 90% IMCSHR with respect to the increase of the number of incremental frames for the color threshold and the similarity color threshold of 0.05. The results in Table 2 show that the matching rate has been kept at a level above 96% with the increase of the number of frames from 1 to 3, which is helpful for dealing with small pose changes, but with more than 3 integration frames, the matching rates are not necessarily better, probably due to segmentation errors.

4.2 *Matching of the Same Moving Objects on Two Different Tracks*

4.2.1 *The same female person on two different tracks (frames 2040 to 2052 versus frames 2150 to 2162)*

The first case reported here is from the same person recorded in two different tracks (track 1, frames 2040 to 2052, and track 2, frames 2150 to 2162, in steps of five frames), and the matching results are shown in Table 3. The

Table 4 Results of IMCSHR matching and post-integration.

Test case	Frame no.	Track no.	Similarity	Matching results
1	0510–0514	1	0.9278	1 (Yes)
	0600–0604	2		
2	0512–0516	1	0.9434	1 (Yes)
	0602–0606	2		
3	0514–0518	1	0.9226	1 (Yes)
	0604–0608	2		
4	0516–0520	1	0.9562	1 (Yes)
	0606–0610	2		
5	0518–0522	1	0.9211	1 (Yes)
	0608–0612	2		
Post-integration	0510–0522	1	N/A	100% (Yes)
	0600–0612	2		

Note: With 90% of major colors cut off, color threshold=0.05, similarity color threshold=0.05, discrimination threshold=0.55, IMCSHR matching threshold=0.8, and final integration matching threshold =0.8.

extracted moving object, moving object mask, and corresponding incremental major color spectrum histogram in typical frames (2048 in track 1 and 2156 in track 2) are shown in Fig. 9.

The results in Table 3 show that in five incremental MCSHR matching test cases, all correctly matched with similarities between 89% and 97% and proved that the proposed IMCSHR algorithm is an accurate major color representation of a moving object and that the matching algorithm proposed in this paper is robust. The post-integration of the five individual matching cases is 1.0, which is higher than the final matching threshold of 0.8, so the two tracks are reliably matched.

4.2.2 *Same blue car in two different tracks (frames 0510 to 0522 versus frames 0600 to 0612)*

The test data reported here are from the same blue car recorded on two different tracks (track 1, frames 0510 to 0522, and track 2, frames 0600 to 0612), and the matching results are shown in Table 4. Two sets of typical extracted moving objects, moving object masks, and their major color histograms are shown in Fig. 10.

The results in Table 4 show us that in five IMCSHR matching test cases, all correctly matched with similarities between 92% and 96%, and the matching rate of the post-integration of the five individual matching cases is 1.0, which is higher than the final matching threshold of 0.8, so the two tracks are also reliably matched.

4.3 *Matching of Two Different Moving Objects on Two Different Tracks*

4.3.1 *A female person in frames 0400 to 0412 versus another female person in frames 2150 to 2162*

The test data reported here are from two different female persons recorded from the same camera but on two different tracks (track 1, frames 0400 to 0412; track 2, frames 2150 to 2162), with two sets of typical extracted moving

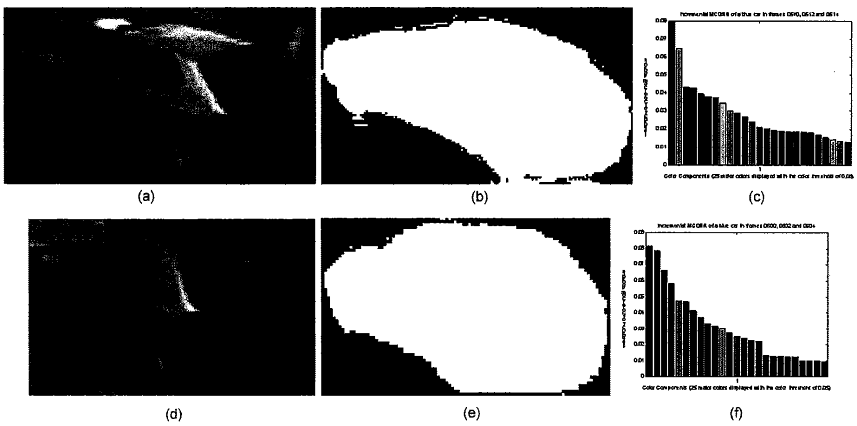


Fig. 10 Moving objects from track 1, frames 0510 to 0522, and track 2, frames 0600 to 0612: (a) MO in frame 0510; (b) mask; (c) IMCSHR of 0510, 0512, and 0514; (d) MO in frame 0600; (e) mask; and (f) IMCSHR of 0600, 0602, and 0604. (Color online only.)

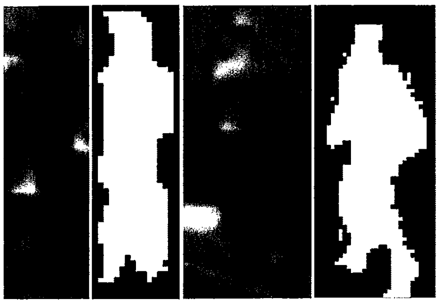


Fig. 11 Moving objects from frames 0400 to 0412 (female person) and frames 2150 to 2162 (different female person).

objects, moving object masks, and their major color histograms shown in Fig. 11. The matching results are shown in Table 5.

The results in Table 5 show us that in all five IMCSHR matching test cases, four cases are correctly discriminated with similarities between 43% and 54%, and the post-integration of the five individual matching cases is 0.2, which is much lower than the final matching threshold of 0.8, so the two tracks are also reliably discriminated. In test case 4, the two moving objects in incremented frames are wrongly matched, and the average similarity in these test cases is higher than the previous test case, probably because the colors of these two female persons are somewhat similar.

4.3.2 *A male person in frames 0850 to 0862 versus a female person in frames 2150 to 2162*

The test data reported here are from two different persons recorded from the same video surveillance camera but on two different tracks (a male person, frames 0850 to 0862, and a female person, frames 2150 to 2162), with two sets of typical extracted moving objects, moving object masks, and their major color histograms shown in Fig. 12. The matching results are shown in Table 6.

The results in Table 6 show that in five IMCSHR matching test cases, all are correctly discriminated with similarities between 14% and 25%, and the post-integration of the five individual matching cases is zero, which is much lower than the final matching threshold of 0.8, so the two tracks are also reliably discriminated. Since the two moving objects have very different colors, they are discriminated easily with the proposed algorithm.

4.3.3 *A blue car in frames 0510 to 0522 versus a white van in frames 0700 to 0712*

The test data reported here are from two different vehicles recorded from the same video surveillance camera but on two different tracks (blue car, frames 0510 to 0522, and white van, frames 0700 to 0712), with two sets of typical extracted moving objects, moving object masks, and their major color histograms shown in Fig. 13. The matching results are shown in Table 7.

Table 5 Results of IMCSHR matching and post-integration.

Test case	Frame no.	Track no.	Similarity	Matching results
1	0400–0406	1	0.5384	0 (No)
	2150–2156	2		
2	0402–0408	1	0.4304	0 (No)
	2152–2158	2		
3	0404–0410	1	0.5410	0 (No)
	2154–2160	2		
4	0406–0412	1	0.9844	1 (Yes)
	2156–2162	2		
5	0408–0414	1	0.5351	0 (No)
	2158–2164	2		
Post-integration	0400–0412	1	N/A	20% (No)
	2150–2162	2		

Note: With 90% of major colors cut off, color threshold=0.05, similarity color threshold=0.05, discrimination threshold=0.55, IMCSHR matching threshold=0.8, and final integration matching threshold =0.8.

Table 6 Results of IMCSHR matching and post-integration.

Test case	Frame no.	Track no.	Similarity	Matching results
1	0850–0854	1	0.1425	0 (No)
	2150–2156	2		
2	0852–0856	1	0.2013	0 (No)
	2152–2158	2		
3	0854–0858	1	0.1985	0 (No)
	2154–2160	2		
4	0856–0860	1	0.1895	0 (No)
	2156–2162	2		
5	0858–0862	1	0.2524	0 (No)
	2158–2164	2		
Post-integration	0850–0862	1	N/A	0% (No)
	2150–2162	2		

Note: With 90% of major colors cut off, color threshold=0.05, similarity color threshold=0.05, discrimination threshold=0.55, IMCSHR matching threshold=0.8, and final integration matching threshold =0.8.

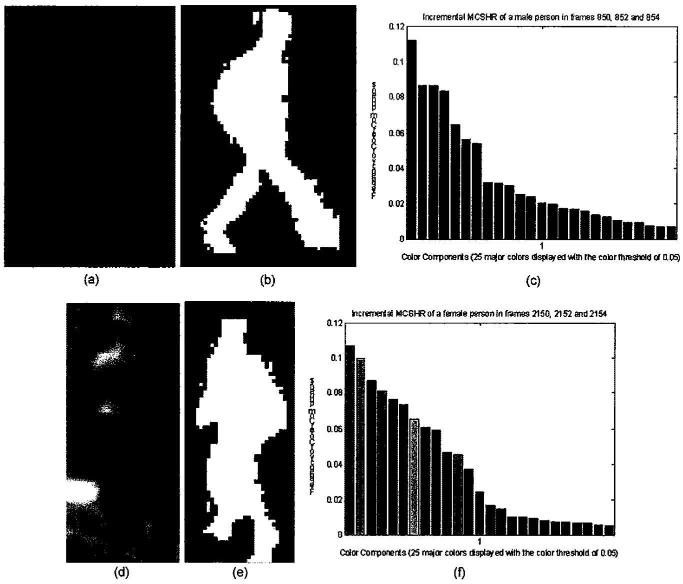


Fig. 12 Moving objects from frames 850 to 862 (male person) and frames 2150 to 2162 (female person): (a) MO in frame 0850; (b) mask; (c) IMCSHR of 0850, 0852, and 0854; (d) MO in frame 2150; (e) mask; (f) IMCSHR of 2150, 2152, and 2154. (Color online only.)

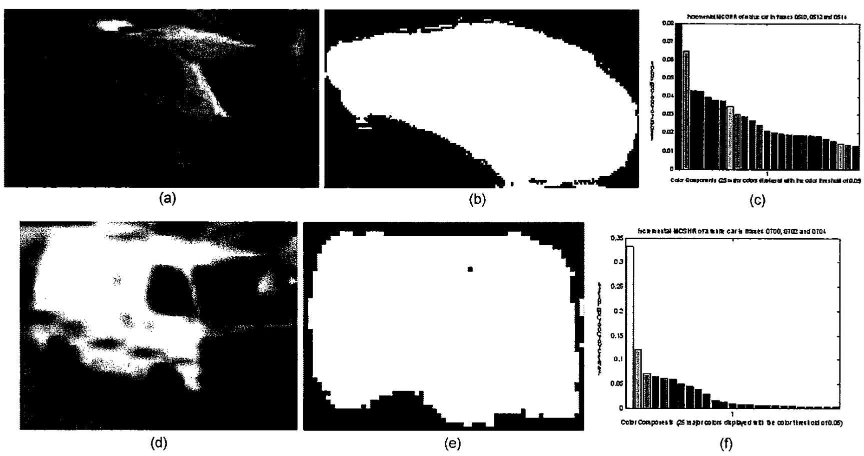


Fig. 13 Moving objects from frames 0510 to 0522 (blue car) and frames 0700 to 0712 (white van): (a) MO in frame 0510; (b) mask; (c) IMCSHR of 0510, 0512, and 0514; (d) MO in frame 0700; (e) mask; and (f) IMCSHR of 0700, 0702, and 0704. (Color online only.)

Table 7 Results of IMCSHR matching and post-integration.

Test case	Frame no.	Track no.	Similarity	Matching results
1	0510–0514	1	0.3297	0 (No)
	0700–0704	2		
2	0512–0516	1	0.2752	0 (No)
	0702–0706	2		
3	0514–0518	1	0.3443	0 (No)
	0704–0708	2		
4	0516–0520	1	0.4119	0 (No)
	0706–0710	2		
5	0518–0522	1	0.3922	0 (No)
	0708–0712	2		
Post-integration	0510–0522	1	N/A	0% (No)
	0700–0712	2		

Note: With 90% of major colors cut off, color threshold=0.05, similarity color threshold=0.05, discrimination threshold=0.55, IMCSHR matching threshold=0.8, and final integration matching threshold =0.8.

The results in Table 7 show that in five IMCSHR matching test cases, all are correctly discriminated with similarities between 28% and 39%, and the post-integration of the five individual matching cases is zero, which is much lower than the final matching threshold of 0.8, so the two tracks are also reliably discriminated. In this test case, the blue car is quite different from the white van, so they are easily discriminated with quite low similarities.

5 Conclusion

In this paper, we present a disjoint track matching algorithm based on incremental major color spectrum histogram matching and post-matching integration algorithms; this algorithm is useful for reconnecting broken tracks due to occlusions and potentially useful for tracking a single object across multiple, disjoint cameras. An incremental major color representation matching algorithm is also used to cope with small pose changes occurring along the track. First an MCSHR is introduced to represent a moving object by a certain percentage (90% in this paper) of its most-frequent colors. Then, a two-directional similarity measurement based on the most-similar-color searching algorithm is used to measure the similarity of the two given moving objects in an IMCSHR in a few frames. Last, our track matching algorithm extends the multiframe matching along the objects' tracks by a post-matching integration algorithm.

Our research shows that the color distance (the value in the range of 0 to 1) defined and used in this paper can accurately measure the similarity between two different colors. The MCSHR algorithm based on the introduced color distance can represent moving objects accurately with the limited number of colors and the frequency of each major color. Experiments have shown that with the increase in the number of integration frames (from 2 to 4), the MCSHR tends to become more stable, in the sense that the actual major colors are emphasized with respect to the other colors. Practically, since there are always errors in the process of moving object detection, together with small pose changes, an integration over a few frames (around 3 or 4) proved helpful for a more accurate and robust moving object representation. Further increase in the number of integrated frames does not necessarily improve the representation accuracy. The experiments have also shown that the matching rate has increased from 93% to 97% with the increase of the number of major colors from 70% to 90% in a given test case, and 90% of major colors is considered as a good trade-off between accuracy and computational load.

Experimental results show that the proposed two-directional similarity measurement based on the most-similar-color searching algorithm can measure the similarity of the two given moving objects accurately—the similarity of two tracks from the same moving objects has proven as high as 89 to 97%, while the similarity of two tracks from different moving objects has been kept as low as 14 to 54%. Experimental results have also shown that with a few (typically 3) frames of MCSHR integration, the proposed IMCSHR algorithm can make matching more robust and reliable than single-frame matching, especially for small pose changes. Based on our experimental experience, the similarity measurement calculation can be saved up to 60% by cutting 10% of the major colors at its tail of representation. Because the post-matching integration is based on the IMCSHR matching results (1 or 0), no detailed feature error will be carried forward to this stage. So, post-matching integration makes track matching more robust and reliable, and it can also avoid false track matching, as long as false IMCSHR matching is less than the track matching threshold.

The proposed track matching algorithms can significantly extend current video surveillance applications by providing them with accurate tracking across nonoverlapping camera views, which is the actual case for many real-world surveillance camera networks. The proposed IMCSHR matching and post-matching integration algorithms have the potential to substantially extend current video surveillance applications. Moreover, to the best of our knowledge, the proposed approach is one of the few in the current literature to tackle the problem of track matching across disjoint tracks and disjoint camera views. Unlike previous papers, our approach does not require global track matching,[18] or rely on a topographic model of the camera network.[19]

Acknowledgments

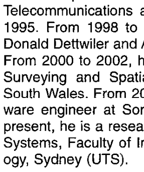
This research is supported by the Australian Research Council, ARC Discovery Grant Scheme 2004 (DP-0452657).

References

1. M. Piccardi and E. D. Cheng, "Multi-frame moving objects track matching based on an incremental major color spectrum histogram matching algorithm," presented at *2nd Joint IEEE International*

Workshop on Object Tracking and Classification in and Beyond the Visible Spectrum, 20 June 2005, San Diego, CA.
2. M. Piccardi and E. D. Cheng, "Track matching over disjoint camera views based on an incremental major color spectrum histogram," presented at *IEEE International Conference on Advanced Video and Signal based Surveillance*, 2005, Como, Italy.
3. E. D. Cheng and M. Piccardi, "Track matching by major color histograms matching and post-matching integration," *13th International Conference on Image Analysis and Processing (ICIAP)*, 6–8 September 2005, Cagliari, Italy, LNCS 3617, pp. 1148–1157 (2005).
4. T. H. Chang and Gong, "Tracking multiple people with a multi-camera system," in *Proc. 2001 IEEE Workshop on Multi-Object Tracking*, pp. 19–26 (2001).
5. I. Haritaoglu, D. Harwood, and L. S. Davis, "W4 real-time surveillance of people and their activities," *IEEE Trans. Pattern Anal. Mach. Intell.* **22**(8), 809–830 (2000).
6. I. Haritaoglu, D. Harwood, and L. S. Davis, "W4: who? when? where? what? a real time system for detection and tracking people," in *Proc. 3rd IEEE International Conference on Automatic Face and Gesture Recognition*, pp. 222–227 (1998).
7. L. M. Fuentes and S. A. Velastin, "People tracking in surveillance applications," in *Proc. 2nd IEEE Workshop on Performance Evaluation of Tracking and Surveillance* (2001).
8. I. Haritaoglu, R. Cutler, D. Harwood, and L. S. Davis, "Backpack: detection of people carrying objects using silhouettes," in *Proc. 7th IEEE International Conference on Computer Vision*, Vol. 1, pp. 102–107 (1999).
9. X. Varona, J. Gonzalez, F. X. Roca, and J. J. Villanueva, "Track: image-based probabilistic tracking of people," in *Proc. 15th International Conference on Pattern Recognition*, Vol. 3, pp. 1110–1113 (2000).
10. H. J. Zhang, J. Wu, D. Zhong, and S. W. Smoliar, "An integrated system for content-based video retrieval and browsing," *Pattern Recogn.* **30**(4), 643–658 (1997).
11. C. Kim and J. N. Hwang, "An integrated scheme for object-based change detection," in *Proc. ACM Int. Multimedia Conf.*, pp. 303–311 (2000).
12. T. Aach, A. Kaup, and R. Mester, "Statistical model-based change detection," *Signal Process.* **31**(2), 165–180 (1993).
13. R. Mech and M. Wollborn, "A noise robust method for segmentation of moving objects in video sequences," in *Proc. ICASSP '97*, Vol. 4, pp. 2657–2660 (1997).
14. A. Neri, S. Colonnese, G. Russo, and P. Talone, "Automatic moving object and background separation," *Signal Process.* **66**, 219–232 (1998).
15. C. Gu and M.-C. Lee, "Semantic segmentation and tracking of semantic video objects," *IEEE Trans. Circuits Syst. Video Technol.* **8**, 572–584 (1998).
16. C. Wren, A. Azarbaygaui, T. Darrell, and A. Pentland, "Pfinder: real-time tracking of the human body," *IEEE Trans. Pattern Anal. Mach. Intell.* **19**(7), 780–785 (1997).
17. A. Lipton, H. Fujiyoshi, and R. Patil, "Moving target classification and tracking from real-time video," in *Proc. IEEE Image Understanding Workshop*, pp. 129–136 (1998).
18. O. Kaved, Z. Rasheed, K. Shafique, and M. Shah, "Tracking across multiple cameras with disjoint views," in *Proc. Ninth IEEE Int. Conf. on Computer Vision*, Vol. 2, pp. 952–957 (2003).
19. D. Makris, T. Ellis, and J. Black, "Bridging the gaps between cameras," in *Proc. 2004 IEEE CS Conf. on Computer Vision and Pattern Recognition*, Vol. 2, pp. 205–210 (2004).
20. Z. Zivkovic and B. Krose, "An EM-like algorithm for color-histogram-based object tracking," in *Proc. IEEE Conference Computer Vision Pattern Recognition* (2004).
21. W. Lu and Y. P. Tan, "A color histogram based people tracking system," in *Proc. IEEE Intl. Symp. Circuits and Systems*, Vol. 2, pp. 137–140 (2001).
22. L. Li, W. Huang I. Y. H. Gu, K. Leman, and Q. Tian, "Principal color representation for tracking persons," in *Proc. IEEE International Conf. on Systems, Man. Cybernetics* (2003).
23. Y. Chen and E. Wong, "Augmented image histogram for image and video similarity search," in *Storage and Retrieval for Image and Video Databases, Proc. SPIE* **3656**, 523–532.
24. J. Hu and A. Mojsolovic, "Optimum color composition matching of images," in *Proc. 15th Int. Conf. on Pattern Recognition*, Vol. 4, pp. 47–51 (2000).

Eric Dahai Cheng received his BE from the Northwest Telecommunications Engineering Institute in 1982, his ME from the National University of Defence Technology in 1986, and his PhD from the University of Sydney in 1999, all in electrical engineering. From 1982 to 1983, he was an assistant radar engineer at Changling Machinery Factory. From 1986 to 1991, he was a lecturer in the Aeronautical Electronic Engineering Department at the Air Force College of Engineering. From 1991 to 1995, he was a lecturer in the Telecommunications Department at the Xian Institute of Posts & Telecommunications and was promoted to associate professor in 1995. From 1998 to 1999, he was a software engineer at MacDonald Dettwiler and Associates Ltd. (Orbital Science Corporation). From 2000 to 2002, he was a research associate at the School of Surveying and Spatial Information Systems, University of New South Wales. From 2002 to 2003, he was a signal processing software engineer at Sonacom System PTY LTD. From 2004 to the present, he is a research associate in the Department of Computer Systems, Faculty of Information Technology, University of Technology, Sydney (UTS).



Massimo Piccardi (MEng 1991, PhD 1995) is an associate professor in the Faculty of Information Technology, University of Technology, Sydney (UTS), where he leads the Computer Vision Research Group. His main research interests are in computer vision, pattern recognition, and video analysis, with applications to video surveillance, human-computer interaction, and multimedia. He has been the author or coauthor of more than 100 scientific papers in international journals and conference proceedings. He has also been the recipient of several competitive research grants and the principal investigator in several research projects, mainly in the area of video surveillance. Dr. Piccardi is a senior member of the IEEE and the IEEE Computer Society and a member of the International Association for Pattern Recognition.