

# Track Matching by Major Color Histograms Matching and Post-matching Integration

Eric Dahai Cheng and Massimo Piccardi

Faculty of Information Technology, University of Technology, Sydney (UTS),  
PO Box 123, Broadway NSW 2007, Australia  
{cheng, massimo}@it.uts.edu.au

**Abstract.** In this paper we present a track matching algorithm based on the “major color” histograms matching and the post-matching integration useful for tracking a single object across multiple, limitedly disjoint cameras. First, the Major Color Spectrum Histogram (MCSH) is introduced to represent a moving object in a single frame by its most frequent colors only. Then, a two-directional similarity measurement based on the MCHS is used to measure the similarity of any two given moving objects in single frames. Finally, our track matching algorithm extends the single-frame matching along the objects’ tracks by a post-matching integration algorithm. Experimental results presented in this paper show the accuracy of the proposed track matching algorithm: the similarity of two tracks from the same moving objects has proved as high as 95%, while the similarity of two tracks from different moving objects has been kept as low as up to 28%. The post-matching integration step proves able to remove detailed errors occurring at the frame level, thus making track matching more robust and reliable.

## 1 Introduction

Being able to track a single object throughout a network of cameras is an important function for effective video surveillance of wide areas [1-7]. However, in most real-world camera networks it is not possible to track a moving object through a continuity of overlapping camera views. Instead, most often the object has to completely exit from the view of a certain camera before it can reappear under the view of a different one. This common scenario is often referred to as disjoint camera views, where observations of a same object are disjoint in time and space to a certain extent. In order to allow tracking in such a scenario, single-camera tracks of a same object must be matched across neighbouring cameras. 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.

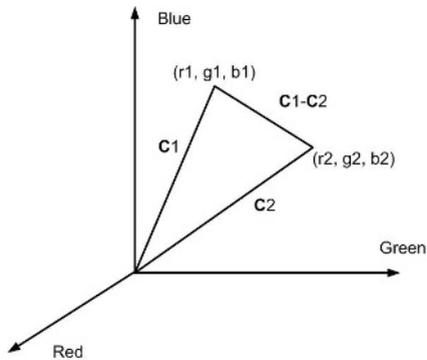
Accordingly, in this paper we present a track matching algorithm based on the “major color” histograms matching and the post-matching integration. First, a color distance based on a geometric distance between two points in the RGB space is introduced to measure the similarity of any two colors. By using the color distance and a given threshold, all pixels from a moving object  $MO_i$  in a given frame  $t$  are clustered into a limited number of 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  $k$  used to represent the moving object. We call this histogram the major color spectrum histogram (MCSH) representation of  $MO_{i,t}$ . Given two arbitrary moving objects,  $MO_{i,t}$  and  $MO_{j,u}$  from two different frames,  $t$  and  $u$ , a similarity criterion based on the major color representation is used to assess their matching (single-frame matching). The single-frame matching is then extended along the two moving objects' tracks by selecting a same number of frames in each track, performing the matching between the corresponding frames of each track, and integrating the matching results along time. Finally, 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 camera views [8, 9]. Differently from those previous papers, our approach does not require global track matching [8] or rely on a topographic model of the camera network [9].

## 2 Major Color Spectrum Histogram

### 2.1 Concept of Color Distance

In this paper, we first introduce a “color distance” between two color pixels in the RGB space based on a normalized geometric distance between the two pixels. Such a geometric distance is defined in equation (1) and exemplified in Fig. 1.



**Fig. 1.** The distance between 2 color pixels in RGB space

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

Where  $C_1$  and  $C_2$  are the color vectors shown in Fig. 1.

### 2.2 Moving Object Major Color Representation

In the RGB color space, using 1 byte to represent each color yields a total of 16.8 million (16,777,216) different colors. It is, in general, very difficult to compare two

objects based on so many possible values. By using the concept of color distance, we can scale down the possible colors from 16.8 million to a very limited number of “Major Colors” (for example, 15 to 100) without losing much accuracy on representing a moving object. For each moving object, a given number of major colors are retained in the representation, while colors that rarely appear are discarded [10,11]. Colors within a given mutual distance threshold are dealt with as a single color. An example of such a major color representation is shown in Fig. 2.

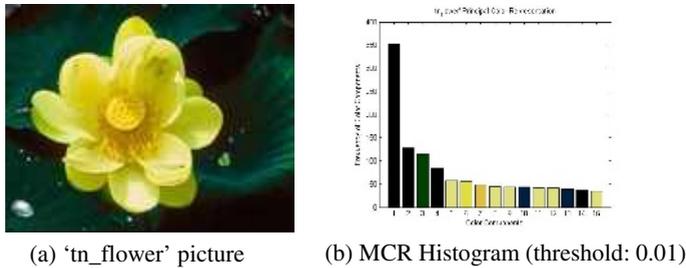


Fig. 2. The Major Color Spectrum Histogram (MCSH) of ‘tn\_flower’

In Fig. 2 (a), the example picture (‘tn\_flower’) is shown. In this picture, we can see that the most frequent colors are around dark green-black and yellow values. Fig. 2 (b) shows us the histogram of the major colors under the color distance threshold of 0.01. In the histogram, we can see that there are 4 main dark green-black bins with the highest frequencies (bins 1-4). The numbers of dark green-black pixels falling in these bins are about 350, 125, 120 and 85 respectively. The yellow colors are distributed in color spectrum bins 5, 6, 7, 8, 9, 11, 12 and 15. The numbers of pixels of yellow colors are between about 60 and 30. There are also 3 dark green-black bins spread in bins 10, 13 and 14, with the pixel numbers between 40 and 35.

### 3 Single-Frame Matching and Post-matching Integration Algorithm

#### 3.1 Moving Objects Similarity

In this paper, the similarity measurement between two moving objects is based on their color histograms [12-14]. In particular, we use the Major Color Spectrum Histogram (MCSH) which is a generalization of [14] in that the objects do not need to be the output of a background subtraction process. We assume that there are  $M$  major

$$MCS(A) = \{C_{A_1}, C_{A_2}, \dots, C_{A_M}\}. \tag{2}$$

colors in the spectrum of moving object A, which can be represented as: Where  $C_{A_i}, i = 1, 2, \dots, M$  is the color vector (RGB) of major colors in object A. Object A’s color spectrum histogram (i.e. the frequencies) can be represented as:

$$p(A) = \{p(A_1), p(A_2), \dots, p(A_M)\}. \quad (3)$$

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

$$MCS(B) = \{C_{B_1}, C_{B_2}, \dots, C_{B_N}\}. \quad (4)$$

Where  $C_{B_j}, j=1,2,\dots,N$  are the color vectors (RGB) of major colors in object B.

Object B's color spectrum histogram can then be represented as:

$$p(B) = \{p(B_1), p(B_2), \dots, p(B_N)\}. \quad (5)$$

In order to define the similarity between two moving objects, a *visibility measurement* of major colors  $C_{A_i}$  from moving object B's major color  $MCS(B)$  is defined as:

$$p(A_i | B) = \min\{p(A_i), \sum_{C_{B_j}: d(C_{A_i}, C_{B_j}) < \sigma} p(B_j)\} \quad (6)$$

with  $i=1,2,\dots,M$ , and  $j=1,2,\dots,N$ . The above equation shows us that the visibility of A from B is given by the sum of histogram values of all major colors in moving object B that are close to the color  $C_{A_i}$  (i.e. the color distance between  $C_{A_i}$  and  $C_{B_j}$  is less than a threshold  $\sigma$ , for example, 0.01, i.e.  $d(C_{A_i}, C_{B_j}) < \sigma$ ). The 'min' operation ensures that  $\sum p(A_i | B) \leq \sum p(A_i)$ . So, the similarity of moving object B to moving object A is defined as:

$$Similarity(B \rightarrow A) = \frac{\sum_{i=1,2,\dots,M} p(A_i | B)}{\sum_{i=1,2,\dots,M} p(A_i)} \quad (7)$$

Similarly, the similarity of moving object A to moving object B is defined as:

$$Similarity(A \rightarrow B) = \frac{\sum_{j=1,2,\dots,N} p(B_j | A)}{\sum_{j=1,2,\dots,N} p(B_j)} \quad (8)$$

where  $p(B_j | A)$  is defined as:

$$p(B_j | A) = \min\{p(B_j), \sum_{C_{A_i}: d(C_{A_i}, C_{B_j}) < \sigma} p(A_i)\}. \quad (9)$$

If both visual objects are the same physical object, both similarities will be high (close to 1.0). Otherwise, both will be low, or at least one will be low (much lower than 1.0). So, we define the overall similarity between moving object A and moving object B as:

$$Similarity(A, B) = \min\{Similarity(A \rightarrow B), Similarity(B \rightarrow A)\} \quad (10)$$

### 3.2 Single-Frame Matching and Post-matching Integration Algorithm

In the track matching algorithm, we consider a same number of frames from each track. The algorithm is shown in Fig. 3.

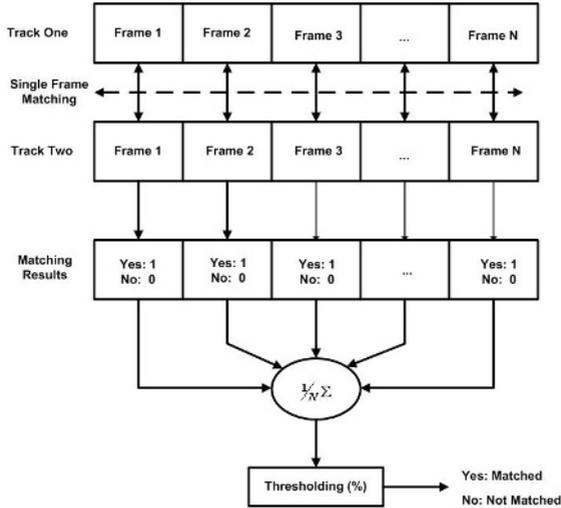


Fig. 3. Single-frame matching and post-matching integration algorithm

Fig. 3 shows us the single-frame matching as the first step of our algorithm. Moving objects from corresponding frames in Track One and Track Two are matched based on similarity of their major color spectrum, and the matching results are given as a binary decision. The second step is the multi-frame post-integration, normalization, and thresholding. The advantages of this algorithm are:

- 1) The single-frame matching is based on the major color spectrum histogram and two direction similarities measurements, which makes the single-frame matching very accurate.
- 2) The final conclusion is made based on the statistical average of single-frame matching. So, no detailed feature errors are carried forward after this stage, which makes the track matching conclusion more reliable than single frame matching.

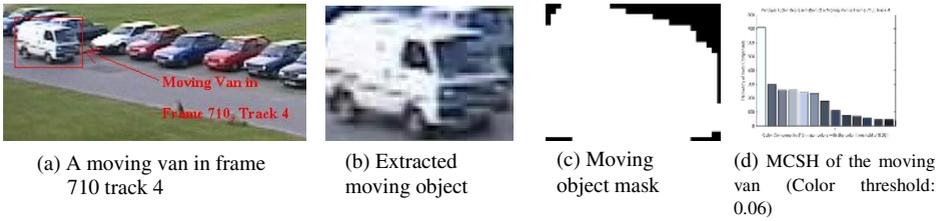
## 4 Experimental Results and Analysis

In our experiments, we segment and track moving objects based on [6, 7]. In the following, we report example results from four typical tracks from the PETS 2001 dataset where three moving objects have been detected and tracked, namely a white van, a female person and a male person. The segmented moving objects, major color spectrum histograms and experimental results are shown in the following sections.

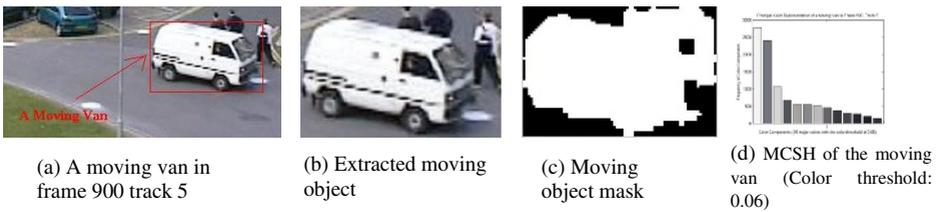
### 4.1 The Matching of the Same Moving Van in Two Different Tracks

The test data here reported are from a white moving van (reference: track 4, frames 700-745, and track 5, frames 900-945.). Typical frames, the extracted moving objects, their masks and major color spectrum histograms are shown in Figs. 4 and 5. (The segmentation was done by a previous researcher in our research group, in which occlusion and shadowing were not taken into account.) The single frame matching

and post-matching integration results with the color distance threshold of 0.01, single frame matching threshold of 0.75, final integration matching threshold of 0.8 are shown in table 1.



**Fig. 4.** Major Color Spectrum Histogram (MCSH) of a moving van (frame 710, track 4)



**Fig. 5.** Major Color Spectrum Histogram (MCSH) of a moving van (frame 900, track 5)

**Table 1.** Results of Single Frame Matching and Post-Integration (Track 4, Frame 700-745/Track 5, Frame 900-945)

Test Case	Frame Index No	Track No	No of colors	Single Frame Similarity	Matching Results(1/0)
1(700_4/900_5)	700	4	110	0.803	1 (Yes)
	900	5	601		
2(705_4/905_5)	705	4	59	0.866	1 (Yes)
	905	5	253		
3(710_4/910_5)	710	4	17	0.805	1 (Yes)
	910	5	191		
4(715_4/915_5)	715	4	29	0.726	0 (No)
	915	5	69		
5(720_4/920_5)	720	4	41	0.946	1 (Yes)
	920	5	54		
6(725_4/925_5)	725	4	22	0.800	1 (Yes)
	925	5	135		
7(730_4/930_5)	730	4	7	0.807	1 (Yes)
	930	5	135		
8(735_4/935_5)	735	4	41	0.809	1 (Yes)
	935	5	148		
9(740_4/940_5)	740	4	29	0.798	1 (Yes)
	940	5	85		
10(745_4/945_5)	745	4	20	0.880	1(Yes)
	945	5	85		
Single Frame Matching Integration	7000-745	4	N/A	N/A	0.9 (90%)
	900-945	5	N/A		

The test results in table 1 show that:

- 1) In most single frame matching test cases, the similarities are between 0.80 to 0.95 which are higher than the single frame matching threshold,  $Th_{sf} = 0.75$  (75%). Thus, the moving objects in these frames are correctly matched. In single frame matching case 4, the similarity is 0.73 and thus matching failed, possibly because of tracking or moving object extraction errors.
- 2) In the example, the post-matching integration rate is 0.9 (90%), which is higher than the final matching threshold  $Th_{track} = 0.8$  (80%); thus, the correct final conclusion can be made.
- 3) Thanks to the accuracy of this procedure, both thresholds  $Th_{sf}$  and  $Th_{track}$  can be kept relatively high so as to strongly limit false positives.

### 4.2 The Matching of Two Different People from Two Different Tracks

The single-frame matching and post-matching integration results for tracks from two different people (reference: moving female person, track 2, frames 460-505; moving male person, track 6, frames 1000-1045) are shown in table 2. Typical frames, the extracted moving objects, their masks and major color spectrum histograms are shown in Figs. 6 and 7.

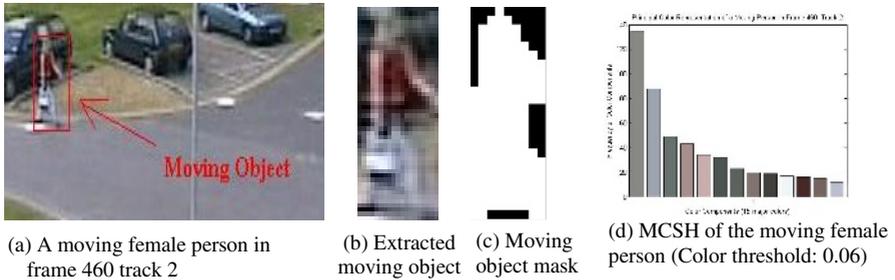


Fig. 6. MCSH of a moving person (frame 460, track 2)

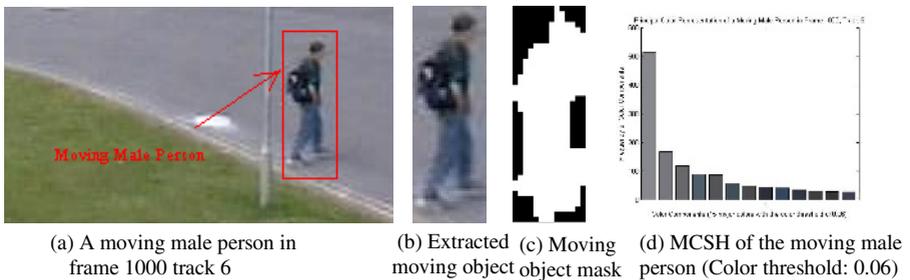


Fig. 7. MCSH of a different moving person (frame 1000, track 6)

The test results in table 2 show that:

- 1) The two moving objects are of similar, small size in their respective frames and their color appearances show some resemblance. However, their MCSH differ significantly. In all test cases, their similarities are between 0 and 0.278 which are well below  $Th_{sf}$ , thus they are correctly discriminated.
- 2) After post-matching integration, the matching rate is still 0 (0%). Therefore, the correct final decision can be easily made.
- 3) Even with the very limited number of major colors used (15), the correct matching was still easily achieved.

The cases described are exemplary of the accuracy of the proposed track matching algorithm. Accuracy is high when views of a same object do not change significantly, while major illumination and pose changes will eventually compromise the effectiveness of matching. In order to make application more general, we are currently developing a compensation algorithm for global illumination variations based on color calibration and background models, and an *incremental* major color spectrum histogram (IMCSH) able to cope with small pose and appearance changes occurring along the track. In addition, the track matching procedure will be eventually integrated with other geometric features such as gait-filtered height.

**Table 2.** Results of Single Frame Matching and Post-Integration (Number of Colors: 15; Color Threshold: 0.01; Track 2, Frame 460-505/Track 6, Frame 1000-1045)

Test Case	Frame Index No	Track No	Single Frame Similarity	Matching Results(1/0)
1 (460_2/1000_6)	460	2	0.0833	0 (No)
	1000	6		
2 (465_2/1005_6)	465	2	0.0372	0 (No)
	1005	6		
3 (470_2/1010_6)	470	2	0.0272	0 (No)
	1010	6		
4 (475_2/1015_6)	475	2	0.2197	0 (No)
	1015	6		
5 (480_2/1020_6)	480	2	0.2779	0 (No)
	1020	6		
6 (485_2/1025_6)	485	2	0.1449	0 (No)
	1025	6		
7 (490_2/1030_6)	490	2	0	0 (No)
	1030	6		
8 (495_2/1035_6)	495	2	0.1912	0 (No)
	1035	6		
9 (500_2/1040_6)	500	2	0.0281	0 (No)
	1040	6		
10 (505_2/1045_6)	505	2	0.0394	0 (No)
	1045	6		
Single Frame Matching Integration	460-505	2	N/A	0 (0%)
	1000-1045	6		

## 5 Conclusions

In this paper, a track matching algorithm has been proposed to match tracks from single objects across non-overlapping camera views. First, a color distance based on a normalized geometric distance between two points in the RGB space is defined and used to measure similarity of two different colors. Then, a Major Color Spectrum Histogram is introduced to represent a moving object by its “major colors” and their frequencies. A similarity measurement is then used to measure the similarity of any two moving objects. Finally, track matching is based on the post-matching integration of single-frame matching.

Based on our experimental results, the following conclusions can be drawn:

- 1) Experimental results shown that the major color spectrum histogram (MCSH) based on the given color distance proved able to represent moving objects accurately with a limited number of colors and their frequencies.
- 2) In the experiments reported, the similarity of a same moving object in two different tracks has reached as high as 95%, while the similarity of two different moving objects has been kept as low as 0% to 28%. This allowed us to use a relatively high threshold (75%) able to limit false positive errors.
- 3) Since the post-matching integration is based on single-frame matching binary results, no detailed feature error is carried forward after this stage. Moreover, post-matching integration makes track matching more robust and reliable than single frame matching.

The proposed track matching algorithm can significantly extend current video surveillance applications by providing them with accurate tracking across non-overlapping camera views which is the actual case for many real-world surveillance camera networks.

## Acknowledgment

This research is supported by the Australian Research Council, ARC Discovery Grant Scheme 2004 (DP0452657).

## References

1. T.H. Chang and S. Gong, "Tracking Multiple People with a Multi-Camera System", Proceedings of the 2001 IEEE Workshop on Multi-Object Tracking, 19-26, 2001.
2. I. Haritaoglu, D. Harwood and L.S. Davis, "W4 Real-Time Surveillance of People and Their Activities", IEEE Trans. on PAMI., 22(8), 809-830, 2000.
3. L.M. Fuentes and S.A. Velastin, "People Tracking in Surveillance Applications", Proceedings of the 2nd IEEE Workshop on Performance Evaluation of Tracking and Surveillance, 2001.
4. C. Wren, A. Azarbaygani, T. Darrell and A. Pentland, "Pfunder: Real-Time Tracking of the Human Body," IEEE Trans. PAMI, 19(7), 780-785, 1997.
5. A. Lipton, H. Fujiyoshi, and R. Patil, "Moving target classification and tracking from real-time video," in Proc. Of the IEEE Image Understanding Workshop, 1998, pp. 129-136.

6. A. Elgammal, D. Harwood, and L. Davis. "Non-parametric Model for Background Subtraction", 6th European Conference on Computer Vision, 2000.
7. A. Senior, A. Hampapur, Y.-L. Tian, L. Brown, S. Pankanti, and R. Bolle, "Appearance Models for Occlusion Handling", 2nd Int. Workshop on Performance Evaluation of Tracking and Surveillance Systems, 2001.
8. O. Kaved, Z. Rasheed, K. Shafique, M. Shah, "Tracking Across Multiple Cameras With Disjoint Views," in Proc. of the Ninth IEEE Int. Conf. on Computer Vision (ICCV'03), vol. 2, pp. 952-957.
9. D. Makris, T. Ellis, and J. Black, "Bridging the Gaps between Cameras," in Proc. of the 2004 IEEE CS Conf. on Computer Vision and Pattern Recognition, vol. 2, pp. 205-210.
10. Zoran Zivkovic and Ben Krose, "An EM-like algorithm for color-histogram-based object tracking," in Proc. IEEE CVPR 2004.
11. W. Lu and Y. P. Tan, "A Color Histogram Based People Tracking System", Proc. IEEE Int'l Symp. Circuits and Systems, vol. 2, pp. 137-140, 2001.
12. Y. Chen and E. Wong, "Augmented image histogram for image and video similarity search," in Proc. SPIE Storage and Retrieval for Image and Video Databases," pp. 523-532.
13. J. Hu and A. Mojsolovic, "Optimum color composition matching of images," in Proc. 15<sup>th</sup> Int. Conf. on Pattern Recognition," vol. 4, pp. 47-51, Barcelona, Spain, 2000.
14. Liyuan Li, Weimin Huang, I.Y.H. Gu, K. Leman, Qi Tian, "Principal Color Representation for Tracking Persons," in Proc. of IEEE Int. Conf. on Systems, Man and Cybernetics 2003, vol. 1, pp. 1007-1012.