

3-2008

Comparing Discrimination and CFA for selecting Tracking Features

Damian M. Lyons
Fordham University

D. Frank Hsu
Fordham University

Follow this and additional works at: https://fordham.bepress.com/frcv_facultypubs

Part of the [Robotics Commons](#)

Recommended Citation

Lyons, Damian M. and Hsu, D. Frank, "Comparing Discrimination and CFA for selecting Tracking Features" (2008). *Faculty Publications*. 22.
https://fordham.bepress.com/frcv_facultypubs/22

This Article is brought to you for free and open access by the Robotics and Computer Vision Laboratory at DigitalResearch@Fordham. It has been accepted for inclusion in Faculty Publications by an authorized administrator of DigitalResearch@Fordham. For more information, please contact considine@fordham.edu.

Comparing Discrimination and CFA for selecting Tracking Features

Damian M. Lyons and D. Frank Hsu,

*Robotics & Computer Vision Laboratory
Department of Computer & Information Science
Fordham University, Bronx, NY 10458, USA
{lyons, hsu}@cis.fordham.edu*

ABSTRACT

The ability of a tracker to isolate the foreground target from the background of an image is crucially dependent on the set of features selected for tracking. Collins & Liu [2] propose an on-line, adaptive approach to selecting the set of features based on the insight that the set of features that best discriminate between target and background classes is the best set to use for tracking. In previous work [10], we have proposed an approach based on Combinatorial Fusion Analysis for selecting features for Real-Time tracking. We discuss the relative merits of the two methods and motivate their combination to produce an improved tracking system. We show several results from a difficult tracking sequence with human targets to demonstrate the effectiveness of the combined system.

Keywords: Target Tracking, Sensor Fusion, Information Fusion, Video Tracking, Surveillance

1. INTRODUCTION

Automated video target tracking has a wide range of applications in surveillance, homeland security, robotics as well as in consumer electronics devices such as cell phones and camcorders. However, there are still some key issues with deploying automated video target tracking in unconstrained scenes, including handling mutual target interactions [5] and interactions between target and background [2]. Interactions between targets include targets occluding one another, partially or completely, and multiple targets appearing to merge together into a single target. Interactions between target and background include targets being occluded by the background and targets appearing to merge into the background due to lighting changes or changes in the background scene. In this paper, we present a novel solution to these two issues based on integrating two existing approaches: combinatorial fusion analysis [6] applied to video target tracking using a rank-score diversity measure [10] developed previously by the authors, and a target background/foreground discrimination measure [2] developed by Collins & Liu. The discrimination approach was developed by Collins & Liu in a framework very similar to combinatorial fusion but was applied by them to single target tracking.

In the following section we briefly overview combinatorial fusion and the rank-score diversity measure, indicate the kind of tracking application where it has been applied and summarize its performance. In Section 3, we describe the discrimination approach used by Collins & Liu and indicate what this adds to the previous work. In Section 4 we describe the integration with our combinatorial fusion target tracker. Section 5 presents results obtained when this approach was evaluated on a difficult tracking sequence, demonstrating the advantage of using both measures. The results show that for a difficult tracking sequence, the combination of both measures can be used to predict which fusion operation to select to produce more accurate multitarget tracking. Finally in Section 6 we summarize our conclusions.

2. TARGET TRACKING USING COMBINATORIAL FUSION

In previous work, the authors have developed a multiple hypothesis based [3] multi-target video tracker that employs combinatorial fusion for improved tracking accuracy [8][9][10][11]. The input to the tracker is a sequence of video images or frames that the tracker processes one by one. In each frame, the tracker determines which regions of the image correspond to potential foreground or target objects and which correspond to the background scene. A data association step matches each foreground region with each current target track hypothesis. A feature similarity measure is used to associate a score, or set of scores (if there are multiple features) with each matching of foreground region to track hypothesis. The growth of the pool of track hypotheses is exponential. Our approach has been to use combinatorial

fusion to address the growth of the hypothesis pool while maintaining tracking accuracy even in cases of multi-target occlusions [9][10].

Previous work in fusion for multisensory video tracking can be divided into three categories [11]: statistical, physical and heuristic. Statistical approaches are the largest category and include Kalman Filtering, JPDA Filtering and more recently, since video tracking rarely abides by Gaussian assumptions, particle-filter based approaches [12]. Physical approaches model the image generation process in sufficient detail to determine how sensory measurements should be fused [11]. In the heuristic approach, a rule of thumb combination operation is used, typically derived from empirical experience, (e.g., [14]). Combinatorial Fusion Analysis (CFA) [6] is an approach to information fusion that has been applied to other fields such as information retrieval, virtual screening and drug discovery, and protein structure prediction. The CFA approach is bottom-up and data-driven. It develops methods and criteria for dynamically selecting *feature subsets and fusion operations* that improve a *performance measure*. Because CFA does not model or make assumptions about target behaviour, it has advantages in hard to model situations such as multiple mutual target occlusions.

In combinatorial fusion analyses applied to target tracking, a set of fusion operators, S , contains the combination operations that can be applied to fuse multiple feature information from each track hypothesis with the objective of calculating a more accurate score for that hypothesis. In [10], S consisted of average rank and Mahalanobis score combinations of the three feature measurements: averaged target color, target bounding-box ratio and target region centroid location. In [9], S consisted of the 11 fusion operations consisting of all combinations of 2 and 3 features fused by average rank combination or Mahalanobis score combination of these same three features. Both [9] and [10] use ground-truth experiments to show that a diversity measure calculated from the rank-score functions of each feature can be used to determine which of the fusion operations to apply to generate improved performance.

Let \mathbf{D} be the set of track hypotheses, $|\mathbf{D}| = n$. Let $s_A(x): \mathbf{D} \rightarrow [0, 1]$ and $r_A(x): \mathbf{D} \rightarrow \{1, 2, \dots, n\}$ be the normalized score function and the rank function of the scoring system for feature A. We define the rank/score function for feature A, $f_A(x): \{1, 2, \dots, n\} \rightarrow [0, 1]$ as:

$$f_A(i) = (s_A \circ r_A^{-1})(i) = s_A(r_A^{-1}(i)) \quad (1)$$

The rank/score diversity measure $d(A, B) = d(f_A, f_B)$ can be defined in several different fashions. In [9][10] we have used the following:

$$d(f_A, f_B) = \sum_{i=1}^n |f_A(i) - f_B(i)| \quad (2)$$

We showed that this diversity could be used to select between rank and score fusions of all three features [10] or, in combination with a relative performance measure [9] to select which pair of three features in addition to rank versus score fusion, for improved tracking accuracy. However, relative performance is only available as a metric in ground-truth experiments, and hence can not be used as part of a general purpose video tracker.

3. TARGET TRACKING USING DISCRIMINATION

Collins and Liu [2] make the point that the performance of a tracker is limited by the features it uses to detect a target. If the features used for tracking cause the target to be appear to be similar to the background (in their case) as well as to other targets (our case), then the tracker performance will suffer. They point out that the best feature for tracking in bright areas might not be the best in shadowed areas for example. They propose that if the selection of color features could be changed from frame to frame, picking the features that best distinguish the target from the background, then the tracking performance would improve.

They select as a color feature space a subset, S_{rgb} , of the linear-weighted score combinations of the red (r), green (g) and blue (b) color channels:

$$S_{rgb} = \{f \mid f = w_r r + w_g g + w_b b \text{ for } w_r, w_g, w_b \in \{-2, -1, 0, 1, 2\}\} \quad (3)$$

The 5^3 features in this space are pruned to 49 by omitting linear combinations, $f_1 = k f_2$, and the case when all coefficients are zero. This reduced feature space S_{rgb}^* includes many commonly used features such as intensity, raw rgb, chrominance features etc.

In our previous work, we have used an averaged, normalized color feature [9] for target tracking. Collins and Liu [2] point out that an average feature is insufficient for background discrimination since it assumes discrimination is a Gaussian problem (see Fig. 2). Instead they construct a histogram of color features in the target foreground region and in the region immediately surround the foreground region (see Fig. 1). The non Gaussian nature of the problem also rules out Linear Discriminant Analysis (LDA) as an approach to separating foreground and background.

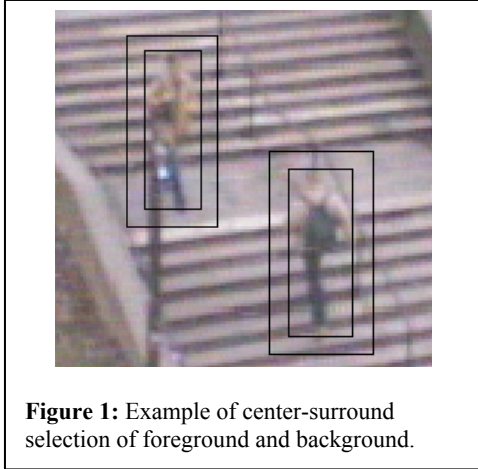


Figure 1: Example of center-surround selection of foreground and background.

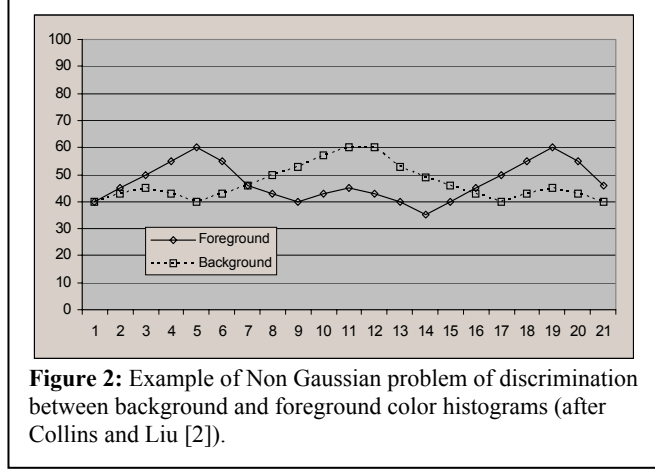


Figure 2: Example of Non Gaussian problem of discrimination between background and foreground color histograms (after Collins and Liu [2]).

Therefore Collins and Liu [2] adopt the approach of evaluating all S_{rgb}^* operations for their target on each frame. Normalized histograms are used to derive discrete probability density functions for each feature in S_{rgb}^* on the target and on the background. The target and background densities are $p: B \rightarrow [0, 1]$ and $q: B \rightarrow [0, 1]$ respectively, where B is the histogram bucket set. They define the log likelihood of a feature value $i \in B$ as

$$L(i) = \log \frac{\max\{p(i), \delta\}}{\max\{q(i), \delta\}} \quad (4)$$

where δ is a small value that prevents division by zero. Positive values of L are associated with the foreground and negative values with the background. The separability of foreground and background under a is defined using the variance of L:

$$var(L; a) = \sum_i a(i) L^2(i) - [\sum_i a(i) L(i)]^2$$

A variance ratio of the total variance over both foreground and background divided by the sum of the variances within foreground and background is used as a criterion to pick the feature that best distinguishes the foreground from background. The variance ratio of the log likelihoods is then defined as:

$$VR(L; p, q) = \frac{var(L; (p + q) / 2)}{var(L; p) + var(L, q)} \quad (5)$$

Collins and Liu using a mean shift tracker operating in each of the N top ranking feature likelihood spaces ($N= 5$ in [2]) computed on the current frame using the densities from the previous frame. The position of the target is obtained by combining the information from the top ranked mean-shift processes. (The means of combination is not described but a JPDA filter approach would seem appropriate.) The approach is shown to work in aerial sequences where targets go from shadowed to bright areas and in sequences where a flag undergoes recurring backlighting effects.

4. INTEGRATION WITH THE CFA TRACKING FRAMEWORK

Our thesis in this paper is that the variance ratio VR when used in conjunction with the rank score diversity d can be used to select a fusion operation to increase tracking accuracy in our combinatorial fusion tracker. The feature space S_{rgb}^* used by Collins and Liu can be interpreted as a fusion operation space on the three color channels; this is the common framework in which the two methods can be combined. However, we will not use the JPDA-like likelihood mean-shift tracking algorithm, we will employ our MHT-based approach, and we will also avoid the combinatorics of a straight extension to multiple targets.

Our CFA tracker only selects fusion operations when the pool of hypothesis reaches resource limits (that is, a threshold size in this paper); hence, the 49 fusion possibilities do not have to be evaluated on every frame for every target. Furthermore, the hypothesis scoring and the rank-score diversity calculations don't need to use the color histogram information, they can contain to use the much simpler average color calculation.

The combined video tracking algorithm has four stages: Determination of foreground regions; Data association; Hypothesis scoring; and, Fusion and pruning. These are described below.

4.1 Determination of Target Regions.

Foreground regions, c_j for $j=1..n_i$, are extracted from each frame i of the image sequence using the non-parametric background estimation technique of Elgammal et al. [4]. A non-parametric distribution is learned for each pixel based on 5 to 10 seconds of video of an initial empty (but not static) scene. Background learning is suspended during target tracking, and pixels are classified as background or foreground based on where they fall with respect to the background distribution learned for that pixel. This approach is effective in filtering background phenomena that result in multimodal pixel value distributions including moving foliage, rain, a small amount of camera vibration and lighting changes. However, if any potential targets are visible during background learning time, it will corrupt the background model and cause potential foreground/background discrimination issues.

4.2 Data Association.

The foreground regions, c_j for $j=1..n_i$, are potential target measurements. For each frame i in the video sequence, a hypothesis generation module associates these measurements with the set of existing track hypothesis \mathbf{D}_i to produce a new pool of all track hypotheses. Any existing track hypothesis which meets the gating criterion for a component c_j is associated with that region, and a new track hypothesis is generated that is the old hypothesis extended by the region. The gating function is that the position of the next component in a track hypothesis, $p(c_j)$ be within a standard deviation of the predicted position p_k for target k :

$$(p_k - p(c_j))^2 < \sigma_k^2$$

In addition to the extension of tracks by new measurements, each region also gives rise to a new track of length 1 initialized to a fixed new track score (to model newly appearing targets), and each track gives rise to a new track of the same length with its score modified by a fixed false alarm score (to model false alarms). The pool of track hypotheses grows as follows:

$$|\mathbf{D}_{i+1}| = |\mathbf{D}_i| \times (n_i + 1) + n_i$$

where n_i is the number of regions segmented from frame i .

4.3 Hypothesis Scoring.

Each track hypothesis in the track pool is given a *raw score* $s = (s_r, s_g, s_b)$ based on the average red, green and blue values of the new region in conjunction with the existing score of the hypothesis as follows. The average color channel values are calculated on the new region:

$$\bar{r} = \frac{1}{N_C} \sum_C r, \quad \bar{g} = \frac{1}{N_C} \sum_C g, \quad \bar{b} = \frac{1}{N_C} \sum_C b$$

where C is the new image region of the target, N_C is the number of pixels in C , and r , g and b are the RGB values of a pixel in the image region. These are compared to the average color values (r_l, g_l, b_l) for the existing track l using

$$sc_r = (r_l - \bar{r})^2, \quad sc_g = (g_l - \bar{g})^2 \quad sc_b = (b_l - \bar{b})^2$$

This similarity value is averaged over all the components (one per frame) of a track hypothesis to obtain the color score for that track hypothesis. If N_l is the existing track length and $(\hat{s}_r, \hat{s}_g, \hat{s}_s)$ the existing raw score, then the new raw score is calculated:

$$s_r = \frac{N_l \hat{s}_r + s c_r}{N_l + 1}, \quad s_g = \frac{N_l \hat{s}_g + s c_g}{N_l + 1}, \quad s_b = \frac{N_l \hat{s}_b + s c_b}{N_l + 1}$$

Note that the similarity value and the score is *smaller* for better hypotheses. The set of target to measurement association hypotheses, including new targets and false alarms, assuming that *at most one measurements matches at most one target*, is then generated and used to calculate a *normalized score* value for each track hypothesis.

For each region c_j , in addition to the average color values, foreground and background color densities, $p_{js}:B \rightarrow [0,1]$ and $q_{js}:B \rightarrow [0,1]$ respectively, where $B=\{1,\dots,32\}$, are generated for each fusion value in $s \in S_{rgb}^*$ in eq. (3).

4.4 Fusion and Pruning.

The pool of track hypotheses grows combinatorially, and needs to be pruned to stay within resource limits. The resource limits are represented by a nominal pool size n_T and the pruning strategy by the rule:

if ($|\mathbf{D}_i| > k_T n_T$) **then** prune \mathbf{D}_i down to size n_T

(The values $n_T=500$, $k_T=1.5$ were used in this paper.) The best performing hypothesis for all targets after fusion were preserved. To find the best performing candidates the pool \mathbf{D}_i was partitioned into a disjoint set of pools D_{ik} for each target k in frame i . For each target pool D_{ik} each of the 49 fusion operations $s \in S_{rgb}^*$ were performed *on the normalized*



Figure 3: Key frames from the sequence used in the ground truth tests.

The sequence is read top to bottom, left to right. There are three targets: one unoccluded target along bottom, and two targets that cross on the left of the image. The two crossing targets are similarly colored, and are similar in color to the background steps making it difficult to track them effectively.

color scores for each hypothesis. For each fusion, two criteria were calculated:

- (1) **Rank/score Diversity:** The hypotheses were scored and ranked according to the three normalized color scores, and the rank/score function for each target for each color score, f_{kr} , f_{kg} and f_{kb} , was generated according to eq (1). The maximum diversity for any pair of the three was then calculated:

$$d_{ks} = \max(d(f_{kr}, f_{kg}), d(f_{kr}, f_{kb}), d(f_{kg}, f_{kb}))$$

In [10] it was shown that a smaller value of rank/score diversity indicated a score fusion produced a better performance, while a larger value indicated that a rank fusion produced a better result.

- (2) **Variance Ratio:** The variance ratio VR_{ks} of the foreground and background log likelihoods was calculated for the terminal region in each hypothesis according to eq. (5). A smaller value of the variance ratio indicates that the background and foreground are spread is large.

The hypotheses for target k that score best on both ratios, that is produces lowest values on both metrics, are considered to be the best performing hypotheses and are kept while the rest are discarded. This process is completed for all targets to bring the pool back to the threshold size of n_T .

5. EXPERIMENTAL RESULTS

To demonstrate the effectiveness of this approach, a ground truth experiment was performed on a difficult tracking sequence, showing that the diversity and variance criteria identify fusion operations with better performance. Key frames from the sequence are shown in Fig. 3. The two crossing targets are of a similar color and similar in color to sections of the background they cross. The sequence was collected at 10 fps and is 10 seconds long. The initial 2 seconds has no targets visible and was used to collect the background. The 5 seconds of the sequence centering on the target cross was used in this experiment. The ground truth for this sequence was constructed by hand on a frame by frame basis and gives the center of each target for each frame.

The tracker was modified so that in the fusion step (Section 4.4) for each target pool D_{ik} the top $m=30$ tracks were evaluated against ground truth using a Mean Sum of Squared Distances (MSSD):

$$MSSD(j) = \frac{1}{nm} \sum_j \sum_i (gp_i - tp_{ij})^2 \quad \text{and} \quad P(j) = \frac{1}{1 + MSSD(j)}$$

where $gp_i, i \in [1, n]$ is the ground truth sequence of target centroid image locations and $tp_{ij}, i \in [1, n]$, is the j th best track's sequence of target centroid image locations. Whichever fusion has the *lower* MSSD measure is the fusion with better performance. Fusions are classified into *positive* and *negative* combinations. A fusion in S_{rgb}^* is considered positive if the tracks it produces are closer to the ground truth than the tracks scored according to the individual features in the fusion. For example, the fusion operation $f = 2r - g$ will be classified as positive, if the tracks it produces have better performance than the tracks produced by either the red (r) or the green (g) color cues alone (based on the hypothesis scoring in section 4.3). We will illustrate that the diversity and variance criteria identify positive fusion operations.

Figure 4 shows a plot of rank/score diversity measure and the variance ration measure against performance for the track hypothesis generated during the ground truth evaluation of the test tracking sequence. The positive and negative fusion cases are shown separately. The negative fusion cases, that is the cases where the fusion produced track hypotheses that did not perform better than the hypotheses of (at least) one of the color channel hypothesis, all tend towards the smaller values of the performance. Smaller values imply a greater difference (larger MSSD values) between the ground truth and the hypotheses – implying less accurate tracking. The positive fusion cases, where the fusion produces better performing track hypothesis, tend towards the larger values of the performance value – implying more accurate tracking. For example, only positive fusion cases for both metrics occur above the performance value of 0.004.

Figure 5 shows a plot of the rank/score diversity against the variance ratio, with positive and negative cases marked separately. Note that the positive cases of each tend to lie towards the lower values of both metrics. A lower value of the variance ratio indicates a larger spread between background and foreground; a lower value of the diversity metric indicates a score fusion will perform well. The plot shows that the better performing fusion operations lie on the lower end of each axis. Therefore, the strategy adopted in the tracker for this paper is to select the fusion with lowest combined value of both rank/score diversity and variance ratio metrics.

Figure 6 shows the end result of the tracking experiment. The top row (Fig. 6 (a) through (c)) in the figure shows the tracks produced for each of the three targets based on using a single color feature, the average RGB value. The top 5

tracks are shown, with higher ranking tracks in lighter shades of gray superimposed on lower ranked tracks. The tracks in the bottom row (Fig. 6 (d) through (f)) show the result of selecting the best fusion operator in S_{rgb} using diversity and variance metrics.

For target 3 there is little effect, since that target underwent no occlusions. All the tracks are very similar and overlay each other, so just the highest ranking tracking is visible.

Figure 6 shows the end result of the tracking experiment. The top row (Fig. 6 (a) through (c)) in the figure shows the tracks produced for each of the three targets based on using a single color feature, the average RGB value. The top 5 tracks are shown, with higher ranking tracks in lighter shades of gray superimposed on lower ranked tracks. The tracks in the bottom row (Fig. 6 (d) through (f)) show the result of selecting the best fusion operator in S_{rgb} using diversity and variance metrics. For target 3 there is little effect, since that target underwent no occlusions. All the tracks are very

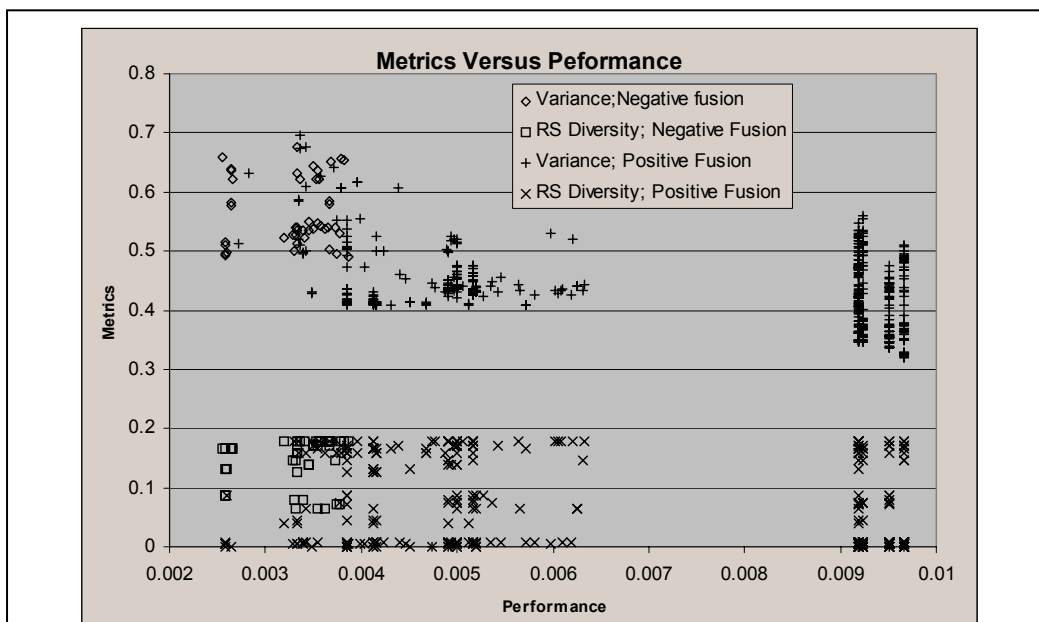


Figure 4: Graph of both Metrics versus Performance.

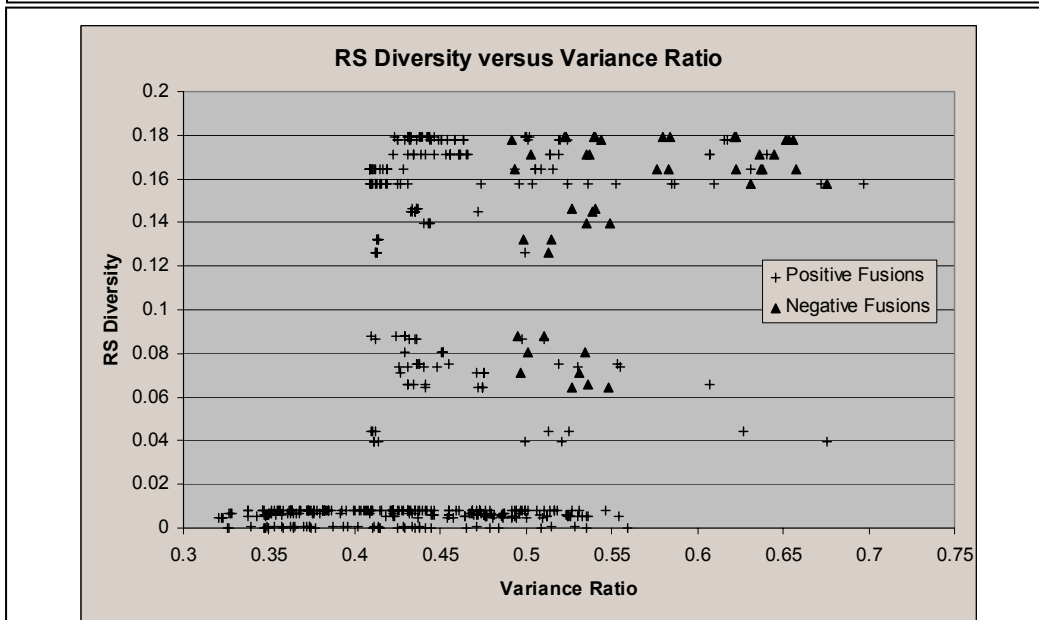
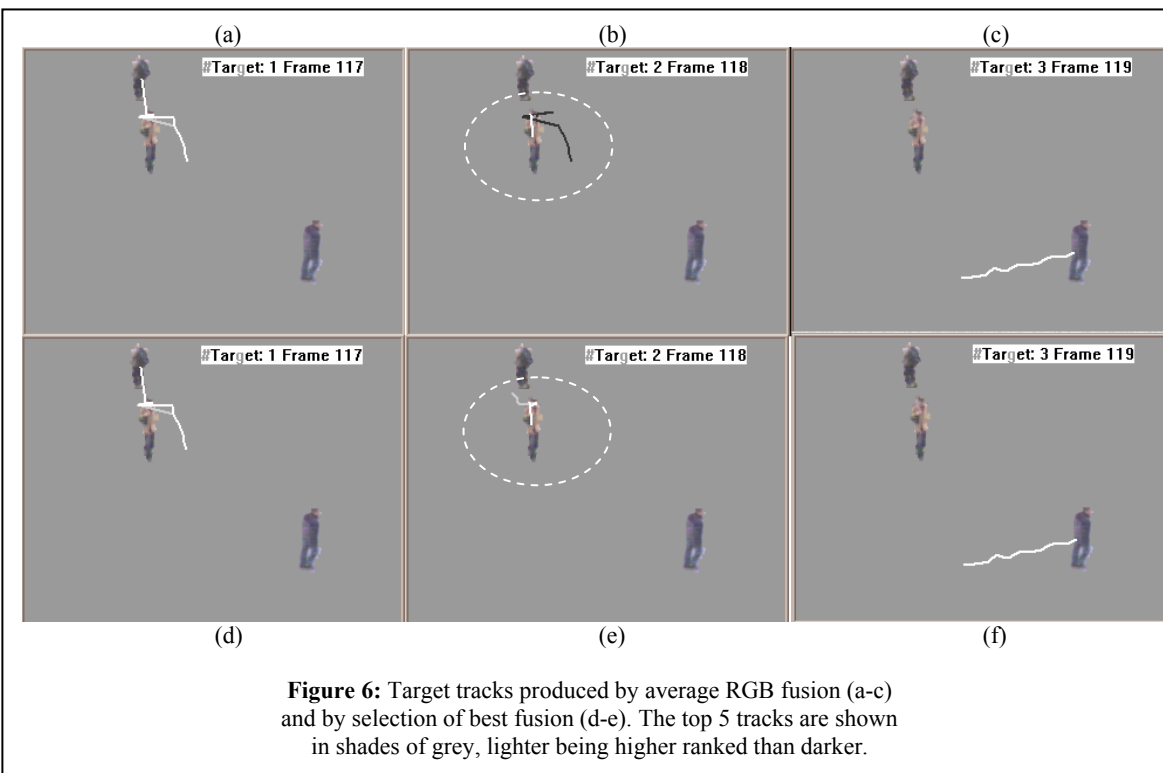


Figure 5: Graph of Rank/Score Diversity versus Variance Ratio



similar and overlay each other, so just the highest ranking tracking is visible.

For targets 1 and 2, which undergo crossing and mutual occlusion (Fig. 3), there is an interference effect: Target 2 loses part of its track before occlusion when the average RGB fusion is used (Fig. 6(b)). Compare the area within the dashed ellipse around target 2 in Fig. 6(b) and Fig. 6(e). One of the lower ranked track hypotheses for target 2 in this case is clearly from target 1 (the track in black in Fig 6(b)). However, selection of the best fusion operation by using the diversity and variance metrics results in the missing portion of the track being preserved (Fig. 6(e)) and the conflation with target 2 is also avoided.

6. SUMMARY

This paper extends previous work in the use of combinatorial fusion analysis for automated video tracking. In previous work in that area [9][10] we have shown that the rank/score diversity measure can be used to select between score fusion and rank fusion operations [10], and when used in conjunction with a relative performance measure [9], to select a rank or score fusion of two features from a set of three features. Here we adopt the color feature space S_{rgb}^* consisting of 49 score fusion operations (weighted linear combinations) of Collins & Liu [2]. We discuss the relative merits of their variance ratio and the rank/score diversity and we present the integration of both into our CFA tracker. Experimental results for a difficult tracking case show that the combination of the variance ratio and rank-score diversity can be used to adaptively select the most accurate fusion operation from S_{rgb}^* .

In previous work on CFA, both score and rank fusion operations were studied. In this paper, only score fusion was included, for compatibility with the Collins & Liu work. However, the non-linear nature of rank fusion operations was shown to have additional value in [10]. The extension of S_{rgb}^* to include rank fusion operations will therefore be investigated.

The track scoring system used in this paper was based on average region color; not a very discriminatory target model. In future work, the foreground densities used here for the variance ratio will be used also in the track hypothesis scoring.

References

- [1] Brown, G., Wyatt, J., Harris, R., and Yao, X.; Diversity Creation Method: A survey and categorization. *Inf. Fusion* 6

- (2005), pp5-20.
- [2] Collins, R., and Liu, Y., On-line selection of discriminative tracking features, IEEE Conf. on Computer Vision. Nice, France (2003), pp. 346-352.
- [3] Cox, I.J. and Hingorani, S.L.; An Efficient Implementation and Evaluation of Reid's Multiple Hypothesis Tracking Algorithm for Visual Tracking. Int. Conf. on Pattern Recognition (1994) pp.437-442.
- [4] Elgammal, A., Harwood, D., and Davis, L.S.; *Nonparametric Model for Background Subtraction*. Proc. 6th European Conference on Computer Vision, (2000), pp.751-767.
- [5] Gabriel, P., Verly, J., Piater, J., Genon, A., The State of the Art in Multiple Object Tracking Under Occlusion in Video Sequences. *Advanced Concepts for Intelligent Vision Systems*, pp. 166–173, 2003.
- [6] Hsu, D.F., Chung, Y.S., and Kristel, B.S.; Combinatorial Fusion Analysis: Methods and Practice of Combining Multiple Scoring Systems. In: (H.H. Hsu, editor) *Advanced Data Mining Technologies in Bioinformatics*, Ideal Group Inc, (2005) in press.
- [7] Hsu, D.F. and Taksa, I., Comparing rank and score combination methods for data fusion in information retrieval, *Information Retrieval* 8 (2005). pp.449-480.
- [8] Hsu, D.F., and Lyons, D.M., A Dynamic Pruning and Feature Selection Strategy for Real-Time Tracking. *19th IEEE Int. Conf. on Adv. Information Networking and Applications*, March 28-30 (2005) pp. 117-124.
- [9] Hsu, D.F., Lyons, D.M., and Ai, J., Feature Selection for Real-Time Tracking. Proceedings of “Multisensor, Multisource Information Fusion: Architectures, Algorithms, and Applications 2006,” at the SPIE Defense and Security Symp., 17-21 April 2006, Orlando (Kissimmee), FL.
- [10] Lyons, D., and Hsu, D.F., Combinatorial Fusion for Target Tracking Using Rank-Score Characteristics. *To appear: Information Fusion 2008*.
- [11] Lyons, D., and Hsu, D.F., Rank-based Multisensory Fusion in Multitarget Video Tracking. *IEEE Intr. Conf. on Advanced Video & Signal-Based Surveillance*. Como, Italy. (2005).
- [12] Maimon, D., Combination of Video based camera trackers using a dynamically adapted particle filter. *Proc. 2nd Int. Conf. on Computer Vision Theory and Applications*, March 8-11, Barcelona Spain, 2007.
- [13] Nandhakumar, N., and Aggarwal, J.K.; *Physics-based Integration of Multiple Sensing Modalities for Scene Interpretation*. Proc. of the IEEE. V85, N1, Jan. (1997). pp.147-163.
- [14] Snidaro, L., Foresti, G., Niu, R., and Varshney, P.; *Sensor Fusion for Video Surveillance*. 7th Int. Conf. on Information Fusion. Stockholm Sweden, (2004) pp.739-746.
- [15] Yang, J.M., Chen, Y.F., Shen, T.W., Kristal, B.S., and Hsu, D.F.; Consensus scoring criteria for improving enrichment in virtual screening. *J. of Chemical Inf. & Mod.* 45 (2005), pp 1134-1146.