

Agent-Based Moving Object Correspondence Using Differential Discriminative Diagnosis

Mahesh Saptharishi

John B. Hampshire II

Pradeep K. Khosla

Department of Electrical and Computer Engineering
Carnegie Mellon University
Pittsburgh, PA 15213
{mahesh | hamps | pkk}@cs.cmu.edu

Abstract

We propose a novel method for temporally and spatially corresponding moving objects by automatically learning the relevance of the objects' appearance features to the task of discrimination. Efficient correspondence is achieved by enforcing temporal consistency of the relevances for a particular object. Relevances are learned using a technique we have termed "differential discriminative diagnosis." An agent is assigned to each moving object in the scene. The agent possesses the basic capability to decide whether or not an object in the scene is the one it represents. Each agent customizes itself to the object by means of differential discriminative diagnosis as the object persists in the scene. We explain this correspondence scheme as applied to the task of corresponding moving people in a surveillance system.

1 Introduction

There has been an increased interest in distributed surveillance systems in recent years [1,2,3,4,5]. The objective is to provide critical information to the human user in real time. A surveillance network of reasonable size produces massive quantities of information. Much of this information is redundant and can inundate a human operator while distracting him or her from information of substance. A distributed surveillance system that can automatically eliminate the redundancy in the information conveyed to the user is invaluable. Changes in the scene, induced by the motions and actions of people and vehicles, are usually the subject of interest in most urban surveillance scenarios. The ideal distributed surveillance system should be able to track all the motions and interactions of objects and raise appropriate flags when information of importance needs to be conveyed to the user. This paper addresses the problem of temporally corresponding moving objects to facilitate a good interpretation of the objects' actions.

The complexity of motions in the environment precludes the use of simple positional correspondence, i.e., correspondence based purely on the positions of moving objects. Positional correspondence also fails when moving objects are relatively large with respect to the field of view of the sensors. In such situations, other features of the moving objects, such as different appearance traits, need to be put to good use for robust correspondence. *How can we select appearance features so as to facilitate good correspondence?* The measure of goodness of the features we choose not only depends on the object in question, but also on other objects in the scene. A globally "good" set of features can be estimated *a priori*, but only a subset of these features might be relevant to the correspondence of a particular object. We pose the estimation of the relevance of globally good features for corresponding a particular object as an on-line learning task. *Differential discriminative diagnosis* provides a systematic method for estimating the relevance of features and checking the temporal consistency of these features for a particular object.

2 Related Work

Much work has been devoted to efficient object correspondence and tracking. Surveillance systems described in [1,2,3,4] deal with the problem of detecting and tracking moving objects. The system described in [1] uses correlation with *dynamic templates* of the object as a method for temporally corresponding it. An IIR filter is used to adapt the *dynamic templates* over time. The system described in [2] uses linear prediction with Kalman filters of the position and size of the moving objects. The algorithm described by Cohen and Medioni in [3] combines the detection and tracking process. They use a graph representation to generate dynamic templates of each moving object. An object's trajectory is determined by choosing an optimal path through the graph and enforcing a temporal coherence constraint. [4] uses correlation of the moving object's silhouette and template matching.

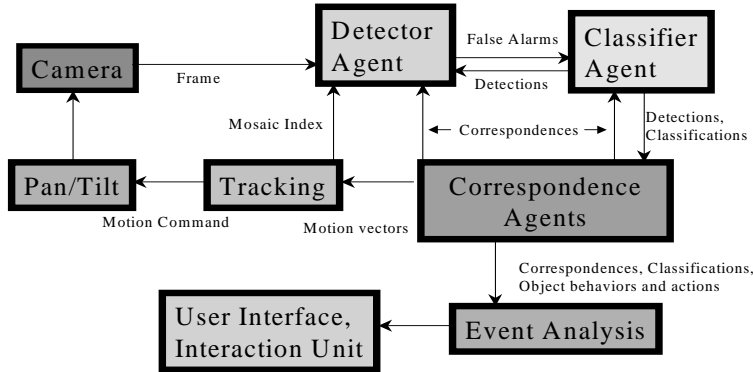


Figure 1. CyberARIES high-level surveillance systems architecture

There has also been a good amount of work done in tracking specific objects. Notably, [6] describes tracking people and their actions. Gaussian models are used to represent 2-D regions or blobs. The model accounts for the position and color of the blobs. These blobs are used to track the position of the person in the scene. Wren and Pentland in [7] extend the notions described in [6] to a 3D context and model a person’s physical actions explicitly. McKenna et al., [8] use a Gaussian mixture model of the color of an object to track it effectively. Black and Jepson in [9] describe an eigenspace method for tracking specific rigid objects. They use a multi-scale eigenspace approach to represent and match objects over time. They apply this technique to the task of tracking and recognizing the gestures of a moving hand. Rehg et al., [10] describe a method for tracking high-DOF articulated objects. They employ this method for tracking humans. They explicitly model the kinematics of articulated parts and use this model to perform correspondence. Other notable people tracking systems include *KidsRoom* [11] and *Cardboard People* [12]. These systems also seek to model the articulation of humans.

Our proposed method relies on knowing the class of the object (person, people or vehicle). Thus, there is domain knowledge incorporated in the correspondence process. The injected domain knowledge not only helps in making the correspondence process robust, but it also helps in making it computationally efficient. By accounting for the different moving objects of interest, we come close to obtaining the versatility of class independent correspondence. In contrast to most of the methods mentioned in this section, our proposed technique poses moving object correspondence as a statistical pattern classification/discrimination problem. Rather than modeling motion, our algorithm finds stable discriminating features to correspond an object. We show that training an agent to correspond an object off-line and giving it the capability to customize itself to the object on-line, leads to an efficient correspondence algorithm.

3 CyberARIES Surveillance Architecture

An agent-based architecture offers an efficient and convenient software infrastructure for a distributed surveillance system. Such an architecture facilitates the combined use of powerful tools from machine learning and computer vision. We have developed an agent based system called CyberARIES for Autonomous Reconnaissance and Intelligent Exploration. CyberARIES has been implemented to run on stationary and mobile surveillance platforms. The object correspondence scheme to be described in this paper uses the CyberARIES architecture as a fundamental implementation tool.

Figure 1 shows the connectivity of the surveillance architecture within CyberARIES with respect to the correspondence agents. The camera produces 320X240 8-bit grayscale images which are sent to the detector agent. The detector agent uses a bank of auto-regressive filters to model the background. The background model is then used to detect moving objects in the scene. It is comparable to the motion detection scheme used in [2]. Connected components analysis is used to segment the detected moving objects from the background. This detection and segmentation scheme has proven to be simple yet effective. The detector agent then feeds a list of segmented objects to the classifier agent which in turn classifies the object as a “person”, “people” or a “vehicle.” The classifier also has the ability to reject detections of no interest to the surveillance task. Examples of uninteresting detections include moving foliage, false alarms caused by changing lighting conditions and high frequency motion of the camera on a vibrating platform. The correspondence agents work with the classified moving objects as their input.

In addition to the feedforward connectivity described so far, there also exist feedback connections from the classifier and the correspondence agents to the detector. The classifier feeds back the locations of the detected objects that were rejected. Information about the predicted future positions of the corresponded moving objects are also fed

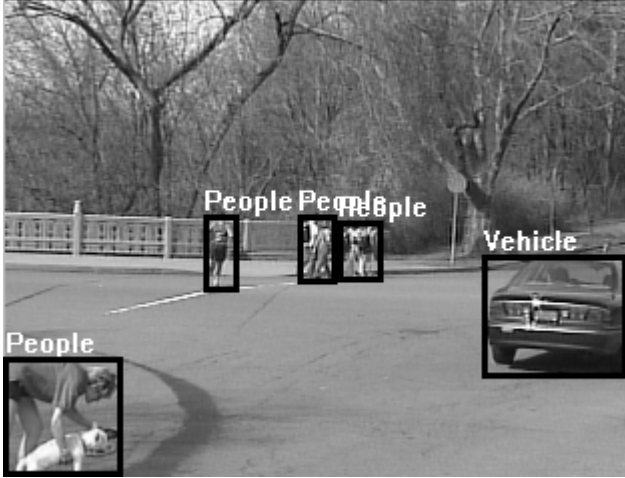


Figure 2. System detects and classifies the people and vehicle

back to the detector. The information feedback is used to adapt the local sensitivity parameters of the detection filters. This simple feedback mechanism is extremely effective in improving the SNR of the detections. Figures 2 and 3 show the system in action with the classifier designed to label each moving object as “people” (in this version of the system, class “people” includes class “person”) or a “vehicle.” The two figures also provide a clear idea of the typical operating environment for the system. Objects can either be a good distance away or very close to the camera.

4 The Basic Correspondence Agent

Consider the detected and classified people and vehicles in Figures 2 and 3. The correspondence agent is responsible for temporally corresponding each moving object. Under the CyberARIES framework, an agent is assigned to every moving object in the scene. This section describes the correspondence algorithm that each agent possesses before any on-line learning occurs.

4.1 Input Representation and the Classification Problem

We pose the correspondence problem as a classification problem. Let $t_{n,s}$ denote an object t at time instance n belonging to a sequence s . The object t is represented by an intensity map with the background subtracted. Figure 4 shows an example of the intensity map contained in $t_{n,s}$. The appearance features that we are interested in are captured by $t_{n,s}$. The temporal correspondence problem can be defined as matching the object $t_{n,s}$ with a previously seen instance $t_{n-1,s}$. Let $X_{n,i,j}$ denote the magnitude of the difference of each pixel between two object

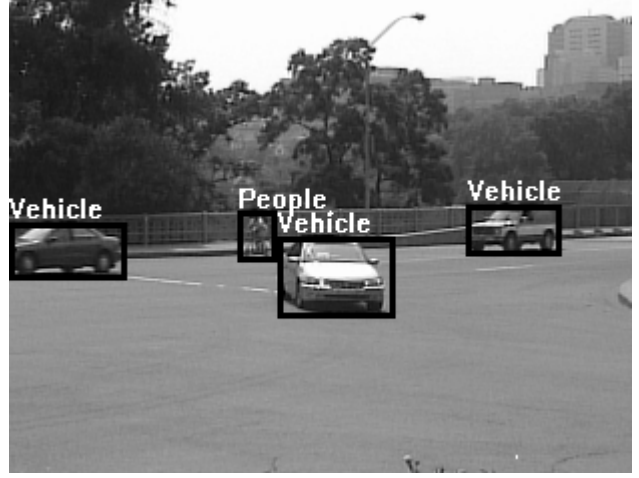


Figure 3. System operating in a slightly different scenario with more vehicles

instances as shown in equation (1). The subscripts i and j

$$X_{n,i,j} = Cr(|Ce(Re(t_{n-1,i})) - Ce(Re(t_{n,j}))|) \quad (1)$$

index potentially different sequences. The function $Re(\)$ resizes the intensity map of the object to a standard size. The function $Ce(\)$ centers the object in the image using its center of mass. The function $Cr(\)$ crops the difference of the intensity maps to a prescribed size. Figure 5 shows the resized, centered and cropped versions of the images in Figure 4. Ideally, if $t_{n,i}$ and $t_{n-1,j}$ represented the same object at two different instances in time, then $X_{n,i,j}$ should contain mostly zeros with very few high magnitude values. Unfortunately, the articulation of a person’s limbs induces large magnitude values, but the locations of these large magnitude values are more or less consistent. Noise and centering errors also cause large magnitude differences. The objective of the classifier is to decide whether or not a given $X_{n,i,j}$ represents the acceptable differences between two instances of the same object or the differences between two instances of different objects. Thus, the classifier classifies each $X_{n,i,j}$ as a “match” or “no match.”

4.2 Designing and Learning a Classifier

An ideal classifier for this application should, with minimum functional complexity, approximate the Bayes-Optimal classifier well. The minimum complexity requirement is critical because of the computational constraints faced by a real-time surveillance system. It should approximate the Bayes-Optimal classifier well because we want the classifier’s performance to be as good as possible given the input representation. Hampshire and Pearlmuter



Figure 4. Two consecutive instances of a moving person from the same sequence. The object $t_{n,s}$ is represented by an intensity map of the person with the background subtracted away as shown above.



Figure 5. The two images shown here are the resized, centered and cropped versions of the images in Figure 4. Equation (1) is the absolute value of the difference between these two images.

[13] prove the equivalence between Multi-Layer Perceptrons (MLPs) and Bayesian discriminant functions for two general classes of objective functions. The two classes can be categorized as error measures and classification figures of merit (CFM) [13]. Hampshire [14,15] shows that *Differential Learning*, using the CFM objective function, generalizes better and requires less functional complexity than error measures such as Mean Square Error. Moreover, *Differential Learning* focuses on maximizing the separation between classes rather than learning the *a posteriori* probabilities of the classes given a finite amount of training data [15].

For this particular application we chose a single output logistic linear neural network trained with *Differential Learning* as the classifier. A total of 249 sequences were available for training the classifier. A total of 120 sequences were used for independent testing. Each sequence contained an average of 15 instances of an object. Sequences were manually sorted from data collections in different environments. Different permutations of sequence pairs were constructed for training and testing. The classifier successfully matched instances of the same moving object with an accuracy of 87%. The 95% confidence interval is [84%, 90%]. Figure 6 shows the weights learned by the classifier. Notice the emphasis of the shoulder and head regions. Emphasis is also placed on the back and sides of the person. The weight layer was forced to be symmetric to account for any bias in the training data for a particular direction of movement for the moving objects.

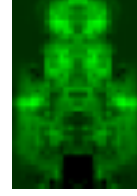


Figure 6. The weights learned by the classifier. Notice the emphasis on the head, shoulders and the sides. The weight layer was forced to be symmetric to account for any bias in the training data

5 Differential Discriminative Diagnosis

The classifier’s training process selects features on the person’s body that help in the classification task given the training data. These features are “globally” relevant, i.e., the selected features help in discriminating a majority of the moving objects without being specific to a particular object. Different environmental conditions and different scenes may reduce or increase the relevance of certain features. More importantly, only a subset of the “globally” relevant features may be applicable to the correspondence of a moving object. In some cases, certain “globally” relevant features may actually hurt the correspondence process. Thus, identifying the feature subset that is relevant to the correspondence of a particular moving object could increase performance dramatically. An agent that represents a moving object customizes itself by estimating the relevance of each feature in the input vector $X_{n,i,j}$ based on the reaction of the classifier to the input and the other objects in the scene. This estimation process is accomplished by means of *differential discriminative diagnosis*.

5.1 Background and Notation

Differential discriminative diagnosis is based on a technique for analyzing the relevance of input features to a classifier in the context of fault detection and diagnosis. This technique was first described in [16]. The relevance analysis method described in [16] is based on two widely used techniques for pruning the weights of neural networks, namely Optimal Brain Damage [17] and Optimal Brain Surgeon [18]. Let $C(X_{n,i,i})$ represent the classifier’s output for the input $X_{n,i,i}$. Let h be defined as in equation (2) where $X_{n,i,j}$ and $X_{n-1,i,i}$ are two distinct input vectors

$$h_{n,i,j} = X_{n,i,j} - X_{n-1,i,i} \quad (2)$$

as defined in equation (1). The output of the classifier for the input $X_{n,i,j}$ can be approximated based on its output for the input $X_{n-1,i,i}$ using the second-order Taylor Series

expansion of the classifier. Equation (3) shows this

$$\begin{aligned} C(X_{n,i,i}) &= C(X_{n-1,i,i} + h_{n,i,j}) \equiv \\ &C(X_{n-1,i,i}) + h_{n,i,j}^T \nabla_X C(X_{n-1,i,i}) \\ &\quad + \frac{1}{2} h_{n,i,j}^T H_X C(X_{n-1,i,i}) h_{n,i,j} \end{aligned} \quad (3)$$

approximation. $H_X C(\cdot)$ represents the Hessian with respect to the input to the classifier, X . We now define a matrix R_i where R_i is a diagonal matrix. Given (2) and (3), let $P_{n,i,j}$ be the difference between the classifier outputs for the inputs $X_{n,i,j}$ and $X_{n-1,i,i}$ as shown in equation (4). If R_i is an identity matrix, then by (3), $P_{n,i,j}$ can

$$P_{n,i,j} = \{C(X_{n-1,i,i} + R_i h_{n,i,j}) - C(X_{n-1,i,i})\} \Big|_{R_i=I} \quad (4)$$

$$\begin{aligned} P &\equiv (R_i h_{n,i,j})^T \nabla_X C(X_{n-1,i,i}) \\ &\quad + \frac{1}{2} (R_i h_{n,i,j})^T H_X C(X_{n-1,i,i}) (R_i h_{n,i,j}) \end{aligned} \quad (5)$$

be approximated as shown in equation (5). If R_i is allowed to deviate from the identity matrix and contains diagonal entries that take on a value other than 1, then each element of $h_{n,i,j}$ is scaled by the corresponding diagonal entry in R_i . This scaling of each dimension of $h_{n,i,j}$ foreshadows the relationship (described in section 5.2) between each entry in R_i and the relevance of the corresponding feature in the feature vector. *Differential discriminative diagnosis* shares conceptual similarities with multi-dimensional scaling, a technique used widely in multivariate data analysis.

5.2 Differential Discriminative Diagnosis

The features that contribute most to the discrimination task are those that consistently appear on the moving object and are different from those on other moving objects. Temporally consistent features have entries with values of zero or a low magnitude in the feature vector $X_{n,i,i}$. We are interested in the discriminative subset of these temporally consistent features that also have a consistently high magnitude in the feature vector $X_{n,i,j}$ where $i \neq j$. We would also like to apply our prior knowledge of feature relevance in the form of the optimized classifier. To this end, we define the *relevance differential* $\mathfrak{R}_{n,i,j}$ for the feature vectors $X_{n,i,i}$, $X_{n,i,j}$ and $X_{n-1,i,i}$ as shown in equation (6). The *relevance differential* is the difference

$$\mathfrak{R}_{n,i,j} = |P_{n,i,j}| - |P_{n,i,i}| \quad (6)$$

between the magnitude of the approximated classifier out-

put differences defined in equation (5). Recall that the index i denotes the sequence of the correct match. The *relevance differential* $\mathfrak{R}_{n,i,j}$, defined in equation (6) indicates the difference between the amassed irrelevances of all the features in the feature vectors $X_{n-1,i,i}$, $X_{n,i,i}$ and $X_{n,i,j}$. Given the feature vectors $X_{n-1,i,i}$, $X_{n,i,i}$ and $X_{n,i,j}$ where $i \neq j$, we find the features that contribute the most to the discrimination task by maximizing the *relevance differential* $\mathfrak{R}_{n,i,j}$ with respect to the matrix R_i . In other words, we wish to make the correct match, $X_{n-1,i,i}$ and $X_{n,i,i}$ as close to each other as possible with respect to the classifier, while making the incorrect match, $X_{n-1,i,i}$ and $X_{n,i,j}$ as far away from each other as possible. Features with a high magnitude entry in R_i are those that are both temporally consistent and are discriminative. Thus, the matrix R_i provides an indication of the relevance of each feature in $X_{n-1,i,i}$ to the correspondence task.

The maximization can be done either by gradient ascent or analytically. The analytical solution is possible because $\mathfrak{R}_{n,i,j}$ as a function of R_i is quadratic with only one local minimum or maximum. Unfortunately, the analytical solution involves computing the inverse of $H_X C(X_{n-1,i,i})$ and the gradient ascent process is too slow for our purpose. For a relatively small logistic linear classifier, the Hessian is ill-conditioned. Thus, even approximating the inverse leaves room for significant approximation errors. In order to make this computation feasible, we choose to assume that the Hessian is diagonal. This assumption doesn't hurt the computation significantly since we can account for the errors in the optimization process. Also, note that the Hessian can be computed off-line except for a multiplicative scalar that depends on the input feature vector $X_{n,i,j}$.

Given the diagonal assumption for the Hessian, we wish to maximize the *relevance differential* $\mathfrak{R}_{n,i,j}$ with respect to R_i for all the moving objects in the scene (indexed by j) and for all time (indexed by n). Thus, the expression to maximize for the i^{th} moving object is given by (7) We refer to equation (7) as the *cumulative relevance*

$$\sum_n \sum_j \mathfrak{R}_{n,i,j} \quad (7)$$

differential, where n iterates over all time that the object was present in the scene, and j iterates through all the other moving objects in the scene at time instance n . We first consider maximizing $\mathfrak{R}_{n,i,j}$ for a particular n , i and j , which maximizes (6). Then we extend the derivation to maximize $\mathfrak{R}_{n,i,j}$ for a fixed n and i over all j . Finally, we derive a recursive equation to maximize over all n , which maximizes (7). Let $^k x$ denote the k^{th} element of a vector x . Also, let M denote the number of elements in the vector x . Similarly, let $^k R_i$ and $^k H_X C(\cdot)$ denote the element at cell (k, k) in each of the matrices. We seek to maximize

equation (6) with respect to each element kR_i of the matrix R_i . Equation (8) represents the value of kR_i at the local

$$S_{n,i,j} = -\text{sgn}(P_{n,i,j}) \times \text{sgn}(P_{n,i})$$

$${}^kR_{n,i,j}^{ext} = \frac{-({}^k h_{n,i,j} + S_{n,i,j} {}^k h_{n,i}) \nabla_X^k C(X_{n-1,i,i})}{({}^k h_{n,i,j}^2 + S_{n,i,j} {}^k h_{n,i}^2) ({}^k H_X C(X_{n-1,i,i}))} \quad (8)$$

extremum of equation (6). Additionally, we define the vectors $A_{n,i,j}$ and $B_{n,i,j}$ of the same size as R_i for notational and computational convenience. Each element, ${}^kA_{n,i,j}$ and ${}^kB_{n,i,j}$, of the vectors $A_{n,i,j}$ and $B_{n,i,j}$ is defined in equations (9) and (10). We can now derive the local extremum of the *relevance differential*, considering all the moving objects (indexed by j) in the scene at time instance n , in equation (11). Finally, equation (11) forms

$${}^kA_{n,i,j} = ({}^k h_{n,i,j} + S_{n,i,j} {}^k h_{n,i}) \nabla_X^k C(X_{n-1,i,i}) \quad (9)$$

$${}^kB_{n,i,j} = ({}^k h_{n,i,j}^2 + S_{n,i,j} {}^k h_{n,i}^2) ({}^k H_X C(X_{n-1,i,i})) \quad (10)$$

$${}^kR_{n,i}^{ext} = \frac{\left(-\sum_j {}^kA_{n,i,j} \right)}{\left(\sum_j {}^kB_{n,i,j} \right)} = -\frac{{}^kA_{i,n}}{{}^kB_{i,n}} \quad (11)$$

$${}^kR_i^{ext} = -\frac{({}^kA_{i,n} + {}^kA_{i,n-1})}{({}^kB_{i,n} + {}^kB_{i,n-1})} \quad (12)$$

$$0 \leq {}^kX_{n-1,i,i} + ({}^kR_i^{max}) ({}^k h_{n,i,j}) \leq 255 \quad (13)$$

the basis for deriving a recursive relationship to find the extremum of equation (7). Equation (12) shows the extremum of the *cumulative relevance differential*. The sign of ${}^kB_{i,n}$ in equation (12) determines if the extremum at ${}^kR_i^{ext}$ is a minimum or a maximum. Based on this fact and the boundary conditions on ${}^kR_i^{max}$ given by equation (13), we can find the ${}^kR_i^{max}$ that maximizes the *cumulative relevance differential* (7).

Each agent adapts its relevance matrix R_i as it sees more instances of the object it represents. The agent can then decide whether or not an object is the one it represents based on the number of relevant pixels (pixels that are temporally consistent and discriminative) on the object. Instead of storing past instances of objects needed to adapt R , the agent stores the sufficient statistics $A_{n,i,j}$

and $B_{n,i,j}$. As a target moves from the field of view of one sensor to another, the agent follows the target.

6 Results

The performance of the correspondence agents was tested on the same 120 independent test sequences used to evaluate the basic correspondence paradigm described in section 4. The agents achieved an accuracy of 96%. The 95% confidence interval is [94.3%, 97.7%]. The customization step shows statistically significant improvements over the 87% accuracy obtained using just the classifier. Figure 7 illustrates the agents powered by *differential discriminative diagnosis*. It shows the relevance of the pixels on the two persons as the agent performs correspondence.

7 Conclusions and Future Work

The proposed correspondence algorithm has been shown to perform well in corresponding people. The algorithm can easily be extended to track vehicles. The algorithm has a few clear failure modes. A temporally abrupt and geometrically drastic change in viewing angle causes the algorithm to fail. The algorithm also fails when tracking a person who bends or twists such that a good number of features that were previously visible are no longer in plain view. We use this correspondence algorithm in conjunction with basic positional correspondence by means of linear prediction of the object's position. This alleviates the effects of some of the failure modes of appearance-based correspondence. We are currently experimenting with situations where more than one sensor is looking at the same target. We hope to extend this algorithm to be able to correspond targets not only within a sensor's field of view, but also among sensors. This raises interesting questions of choosing viewpoint-independent features or multiple sets of viewpoint-dependent features for efficient correspondence.

Acknowledgment

We gratefully acknowledge funding provided for this research under DARPA grant F04701-97-C-0022.

References

- [1] A.J. Lipton, H. Fujiyoshi and R.S. Patil. Moving target classification and tracking from real time video. In IEEE Workshop on Applications of Computer Vision, pp. 8-14, 1998.
- [2] W.E.L. Grimson, L. Lee, R. Romano, and C. Stauffer. Using adaptive tracking to classify and monitor activities in a site. In IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 22-31, 1998.
- [3] I. Cohen and G. Medioni. Detecting and tracking moving objects for video surveillance. In IEEE Computer Society

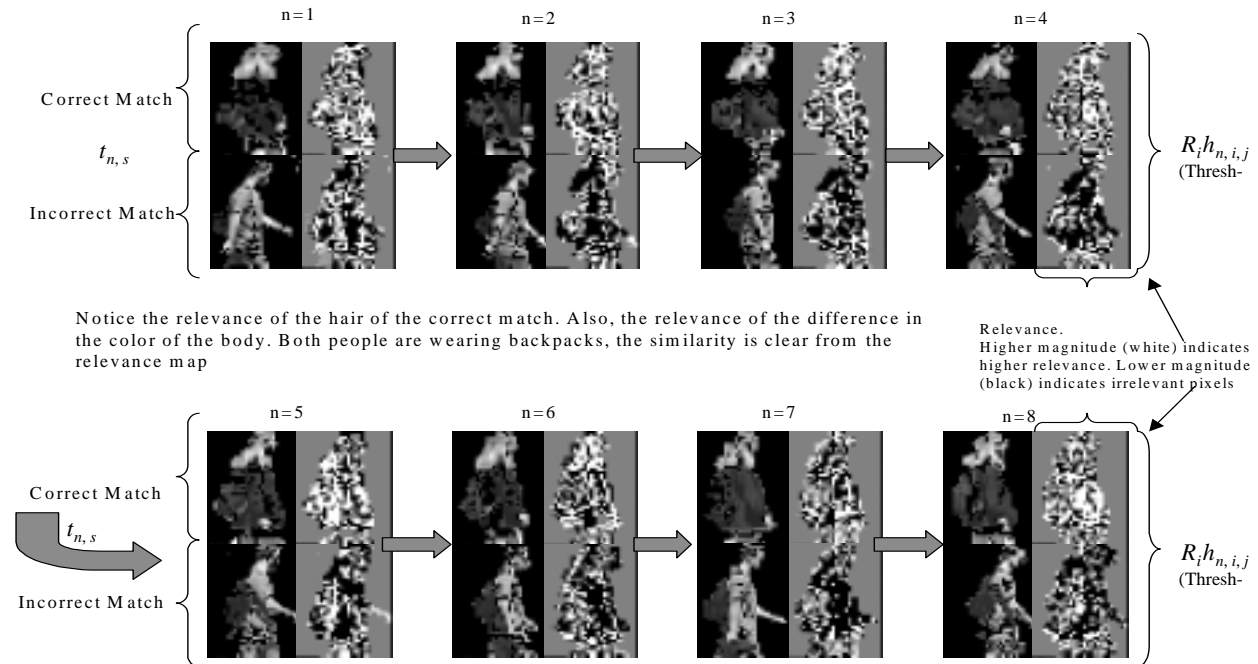


Figure 7. The relevant differences ($R_i h_{n,i,j}$) are shown here for a person being corresponded and another person in the scene. R_h has been thresholded to show the contrast between relevant and irrelevant pixels. The incorrect match clearly has fewer relevant pixels than the correct match.

- Conference on Computer Vision and Pattern Recognition, pp. 319-325, 1999.
- [4] I. Haritaoglu, D. Harwood and L. Davis. W^4 : Who? When? Where? What? A real time system for detecting and tracking people. In IEEE International Conference on Automatic Face and Gesture Recognition, pp. 222-227, 1998.
- [5] C. Diehl, M. Sapharishi, J. Hampshire, and P. Khosla. Collaborative surveillance using both fixed and mobile unattended ground sensor platforms. In SPIE Proceedings on Unattended Ground Sensor Technologies and Applications, vol. 3713, pp. 178-185, 1999.
- [6] C. Wren, A. Azarbayejani, T. Darrell and A. Pentland. Pfnder: real-time tracking of the human body. In IEEE International Conference on Automatic Face and Gesture Recognition, pp. 51-56, 1996.
- [7] C. Wren and A. Pentland. Dynamic models of human motion. In IEEE International Conference on Automatic Face and Gesture Recognition, pp. 22-27, 1998.
- [8] S.J. McKenna, Y. Raja and S. Gong. Tracking colour objects using adaptive mixture models. In Image and Vision Computing 17, pp. 225-231, 1999.
- [9] M.J. Black, A.D. Jepson. EigenTracking: Robust matching and tracking of articulated objects using a view-based representation. In International Journal of Computer Vision, vol.26, no 1, pp. 63-84, 1998.
- [10] J. Rehg and T. Kanade. Model-based tracking of self-occluding articulated objects. In Fifth International Conference on Computer Vision, pp. 612-617, 1995.
- [11] A. Bobick, J. Davis, S. Intille, F. Baird, L.Campbell, Y. Irinov, C. Pinhanez and A. Wilson. Kidsroom: Action recognition in an interactive story environment. In M.I.T. TR No: 398, 1996.
- [12] S. Ju, M.J. Black, Y. Yacoob. Cardboard People: A parameterized model of articulated image motion. In IEEE International Conference on Automatic Face and Gesture Recognition, pp. 38-44, 1996.
- [13] J.B. Hampshire II and B.A. Pearlmutter. Equivalence Proofs for Multi-Layer Perceptron Classifiers and the Bayesian Discriminant Function. In Proceedings of the 1990 Connectionist Models Summer School, pp. 159-172, 1991.
- [14] J.B. Hampshire II and A.Waibel. A Novel Objective Function for Improved Phoneme Recognition using Time-Delay Neural Networks. In IEEE Transactions on Neural Networks, vol. 1, no. 2, pp. 216-228, 1990.
- [15] J.B. Hampshire II. A Differential Theory of Learning for Efficient Statistical Pattern Recognition. Ph.D.Thesis, Carnegie Mellon University, 1993.
- [16] J.B. Hampshire II and D.A. Watola. Diagnosing and Correcting System Anomalies with a Robust Classifier. In IEEE Proceedings of the 1996 International Conference on Acoustics, Speech, and Signal Processing, vol. 6, pp. 3507-3509, May, 1996.
- [17] Y. LeCun, J. Denker and S. Solla. Optimal Brain Damage. In Advances in Neural Information Processing Systems, vol. 2, pp. 598-605, 1990.
- [18] B. Hassibi and D.G. Stork. Second-order derivatives for network pruning: Optimal brain surgeon. In Advances in Neural Information Processing Systems, vol. 5, pp. 164-171, 1993.