

A feature-selection approach to fusing electro-optical and IR imagery

Vinay Sharma and James W. Davis

Exploiting regularities in object shape for feature-level image fusion leads to improved object segmentation.

The demand for persistent 24/7 video surveillance is driving increased deployment of IR-imaging sensors. Since these devices are impervious to ambient light conditions, combining them with traditional electro-optical (visible light) imagers can potentially lead to systems that are effective during day and night. The main challenge is to fuse the complementary and redundant information across both types of sensors in a way that improves the capability and robustness of the surveillance system.

Addressing this challenge, we have developed a technique that combines (fuses) data from registered IR and visible sensors to improve the extraction of object shapes (object segmentation) from surveillance images. While typical fusion algorithms merely gather information into a single image to improve the human visual perception of the scene, our approach focuses on the specific goal of enhancing automatic object segmentation. This is important since object silhouettes are commonly used by computer-vision algorithms for higher-level functions such as object recognition.

We have created a novel solution that tackles fusion essentially as a feature-selection problem.¹ Typically, such an approach is used to identify a set of features that best solves a given classification task.² In contrast, we use this technique to select image features based on a criterion that exploits the regularities observed in object shapes. Our method is applicable to any combination of imaging sensors, provided they are colocated and registered. The processing pipeline and the main computation stages of the algorithm are shown in the flowchart in Figure 1. The feature-extraction stage is preceded by an object-segmentation routine employed in just one of the input sensors (denoted *A*). The routine is used only to bootstrap the feature-selection process. Hence, any method that provides even rough,

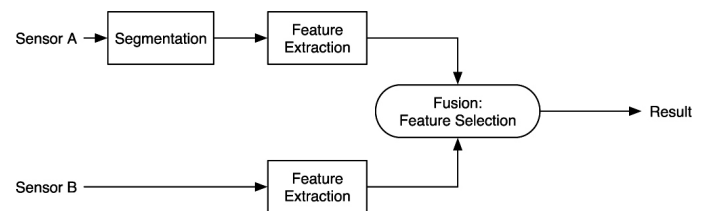


Figure 1. Flowchart of fusion method.

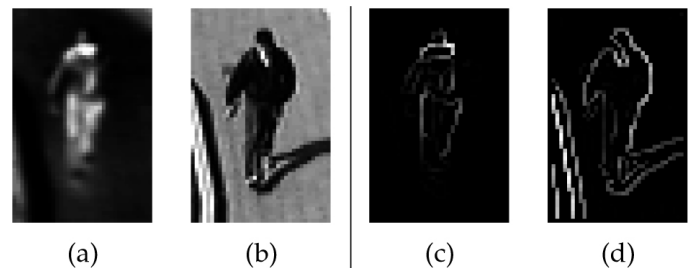


Figure 2. Representative input. (a) IR subimage (sensor A). (b) Visible subimage (sensor B). (c) Object contours obtained from initial segmentation of (a). (d) Edges (showing relative edge strength) obtained from (b).

incomplete object segmentation can be employed at this stage. Contour (edge) features are then extracted from the rough results of sensor *A* and from the corresponding image region of sensor *B*. Figure 2 shows a typical input to the algorithm.

Next, we select features from *B* that are relevant to those in *A*. In the context of fusion, we propose that a set of features has high relevance to another if it provides both redundant and complementary information. The choice of contour features enables us to further define relevance as the ability of a set of features to coincide with and complete object boundaries that have been only partially captured by another set.

Continued on next page

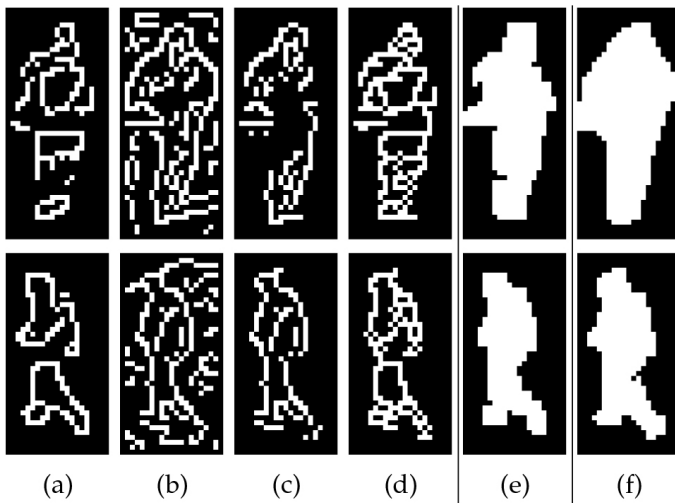


Figure 3. Example fusion results. (a) Contours detected in the IR domain (sensor A). (b) Contours present in the visible domain (sensor B). (c) Contours automatically selected from (b). (d) Overlay of contours from (c) on (a). (e) Segmentation obtained after completing and filling (d). (f) Manually segmented object regions.

We use mutual information as a measure of the relevance of features across the two sensors. The mutual information of two random variables captures to what extent knowledge of one variable reduces the uncertainty in the other. To estimate the probability distribution required for computing this quantity, we rely on the regularities in shape and form found in most objects of interest. We extend the notion of affinity—originally defined to measure the smoothness of the curve joining two edge elements³—to our contour features. Using this new affinity measure,⁴ we formulate conditional probability distributions of the features from sensor A with respect to B. We then compute the mutual information based on the conditional distributions.

The next step is to identify the set of contour features from B that maximizes the mutual information with the set from A. We employ a greedy incremental search scheme that adds to the set of selected features one at a time. The contours from A overlaid on those from B form the fused result, which can then be completed and filled to create silhouettes.

To test our approach, we employed a pair of registered long-wave IR and electro-optical sensors, with the IR camera set as the ‘reference’ device (A). Our experiments¹ showed that the fusion technique generates significantly better object silhouettes than obtained using a sophisticated object-segmentation routine⁵ in the IR domain alone. We also examined how the quality of the initial bootstrap segmentation affects the final result. We found that the proposed approach may start from highly impoverished

information and yet generate overall silhouettes of much higher quality. In fact, our algorithm can maintain reasonable quality even as the required bootstrap segmentation is made so weak as to cover only half of the actual object region. Figure 3 shows example fusion results.

Conclusion

We have developed a new, goal-oriented, feature-level fusion technique for object segmentation. The algorithm treats fusion as a feature-selection problem. The approach exploits the natural structure of the world within a mutual-information framework to define a suitable selection criterion. Our experiments using a pair of IR and electro-optical sensors show that the technique can improve object-segmentation performance over using either sensor alone. In the future, we would like to extend the method to enable two-way information flow in our fusion pipeline. Such an approach would potentially allow the final segmentation to be built up incrementally, such that in each iteration the segmentation from one sensor would seed feature selection in the other, and so on.

Author Information

Vinay Sharma

Digital Signal Processing Research and Development Center
Texas Instruments Inc.
Dallas, TX

Vinay Sharma is a member of the Embedded Vision branch. He received his PhD in computer science and engineering from Ohio State University in 2008, with a major in computer vision. He received a BE (Hons) degree in computer science from the Birla Institute of Technology and Science in Pilani (India) in 2002.

James W. Davis

Department of Computer Science and Engineering
Ohio State University
Columbus, OH
<http://www.cse.ohio-state.edu/~jwdavis/>

James W. Davis is an associate professor of computer science and engineering. He received his PhD from the Massachusetts Institute of Technology in 2000. His research specializes in computer-vision approaches to video surveillance and human-activity analysis.

Continued on next page

References

1. V. Sharma and J. Davis, *Feature-level fusion for object segmentation using mutual information*, in R. Hammoud ed., **Augmented Vision Perception in Infrared**, **Adv. Pattern Recognit.** **26**, ch. 13, Springer, 2009.
2. H. Peng, F. Long, and C. Ding, *Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy*, **IEEE Trans. Pattern Anal. Mach. Intell.** **27**, pp. 1226–1238, 2005.
3. L. Williams and D. Jacobs, *Stochastic completion fields: a neural model of illusory contour shape and salience*, **Neural Comput.** **9** (4), pp. 837–858, 1997.
4. V. Sharma and J. Davis, *Integrating appearance and motion cues for simultaneous detection and segmentation of pedestrians*, **IEEE Int'l Conf. Comput. Vision**, 2007.
5. J. Davis and V. Sharma, *Background-subtraction in thermal imagery using contour saliency*, **Int'l J. Comput. Vision** **71** (2), pp. 161–181, 2007.