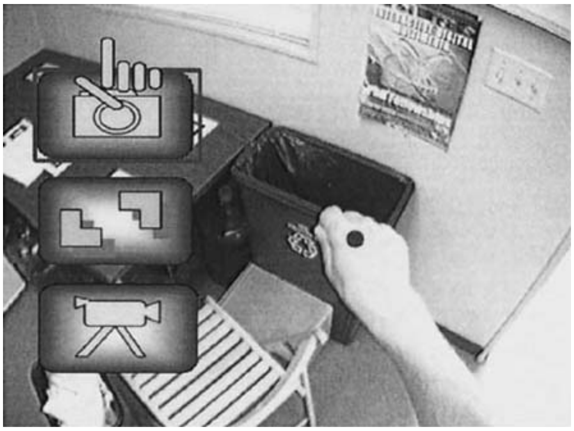
Vision-Based Interaction

Definition

Vision-based human–computer interaction provides a wider and more expressive range of input capabilities by using computer vision techniques to process sensor data from one or more cameras in real-time, in order to reliably estimate relevant visual information about the user.

Human–computer interaction involves information flow in both directions between computers and humans, which may be referred to as *input* (human to computer) and *output* (computer to human). Traditional computer interfaces have very limited input capabilities, typically restricted to keyboard typing and mouse manipulations (pointing, selecting, dragging, etc.). The area of *vision-based interaction* [1] seeks to provide a wider and more expressive range of input capabilities by using computer vision techniques to process sensor data from one or more cameras in real-time, in order to reliably estimate relevant visual information about the user – i.e., to use vision as a passive, non-intrusive, non-contact input modality for human–computer interaction.

In human-to-human interaction, vision is used to instantly determine a number of salient facts and features about one another such as location, identity, age, facial expression, focus of attention, posture, gestures, and general activity. These visual cues affect the content and flow of conversation, and they impart contextual information that is different from, but related to, other interaction modalities. For example, a gesture or facial expression may be intended as a signal of understanding, or the direction of gaze may disambiguate the object referred to in speech as "this" or the direction "over there." The visual channel is thus both co-expressive and complementary to other communication channels such as speech. Visual information integrated with other input modalities (including keyboard and mouse) can enable a rich user experience and a more effective and efficient interaction. Vision-based interaction may be useful in a wide range of computing scenarios in additional to standard desktop computing, especially mobile, immersive, and ubiquitous computing environments. A nice example of simple vision technology used effectively in an interactive environment was the KidsRoom project at the MIT Media Lab [2]. Another example is HandVu [3], which allows users of mobile augmented reality systems to use their hands to drive the interface, by robustly tracking hands and looking for a few known hand gestures/postures. Figure 1 shows HandVu at work.



Vision-Based Interaction. Figure 1. The HandVu system in action, providing hand-based control in a mobile augmented reality system.

In order to provide this kind of input about users, many researchers in the field of computer vision have focused on modeling, recognizing, and interpreting human behavior. Among the most studied sub-areas are face detection and location, face recognition, head and face tracking, facial expression analysis, eye gaze tracking, articulated body tracking, hand tracking, and the recognition of postures, gaits, gestures, and specific activities. Several of these have applications in areas such as security and surveillance, biometrics, and multimedia databases, as well as in human–computer interaction. Although many significant technical challenges remain, there has been notable progress in these areas during the past decade, and some commercial systems have begun to appear. In general, further research needs to improve the robustness and speed of these systems, and there needs to be a deeper understanding of how visual information is best utilized in human–computer interaction.

Cross-References

▶ Human Computer Interactions

References

1. M. Turk and M. Kölsch, "Perceptual Interfaces," in G. Medioni and S.B. Kang (Eds.), "Emerging Topics in Computer Vision," Prentice-Hall, Englewood, Cliffs, NJ, 2004.

2. A. Bobick, S. Intille, J. Davis, F. Baird, C. Pinhanez, L. Campbell, Y. Ivanov, A. Schtte, and A. Wilson, "The KidsRoom: A Perceptually-Based Interactive and Immersive Story Environment," PRESENCE: Teleoperators and Virtual Environments, Vol. 8, No. 4, August 1999, pp. 367–391.
3. M. Kölsch, M. Turk, and T. Höllerer, "Vision-Based Interfaces for Mobility," Proceedings of the MobiQuitous'04, 1st IEEE International Conference on Mobile and Ubiquitous Systems: Networking and Services, Boston, MA, pp. 86–94, August 2004.

Visual Cryptography

Definition

Visual cryptography or visual secret sharing represents a group of effective schemes for image and video hading and watermarking.

Visual cryptography (VC) or visual secret sharing (VSS) schemes [1] constitute probably the most cost-effective solution within a (k, n)-threshold framework. The VSS schemes use the frosted/transparent representation of the shares and the properties of the human visual system to force the recognition of a secret message from overlapping shares without additional computations or any knowledge of cryptographic keys [1–3].

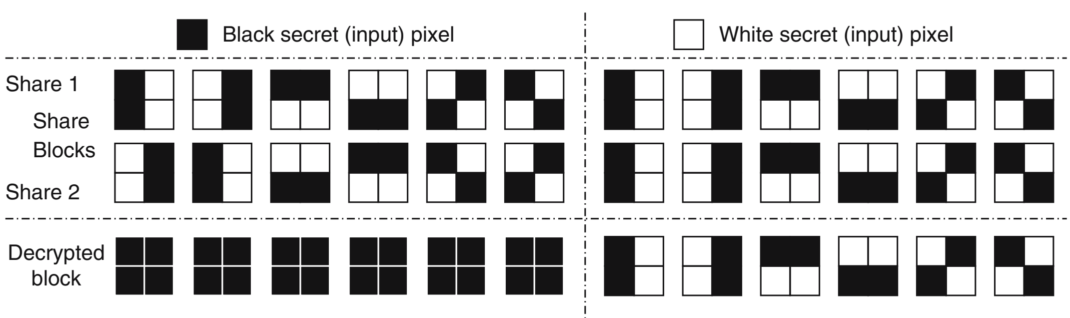
As it is shown in Fig. 1, the conventional VSS schemes operate on a binary input. Following the encryption procedure in a (k, n)-threshold framework, the secret binary pixel is encrypted into n blocks of $m_1 \times m_2$ binary pixels. The actual share blocks are randomly generated through the column permutation of the $n \times m_1 m_2$ basis matrices. By repeating the process for all pixels of a $K_1 \times K_2$ secret (input) image, the VSS scheme produce n binary shares with dimensions of $m_1 K_1 \times m_2 K_2$ pixels. Each noise-like binary share is distributed to one of n participants. The secret image is visually revealed (Fig. 2) only if any k (or more) recipients stack their shares printed as transparencies together on an overhead projector or specialized display.

Cross-References

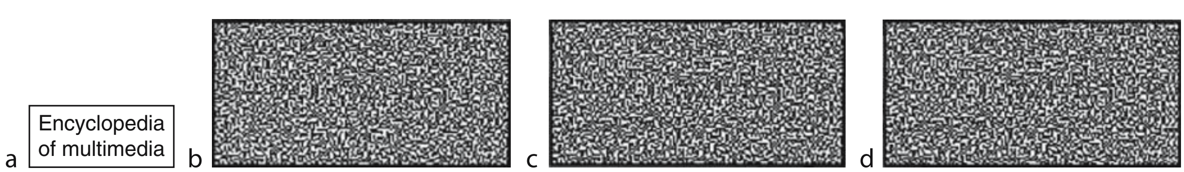
- ► Halftoning Based VSS
- ► Image Secret Sharing
- ► Image Watermarking Using Visual Cryptography
- ► Threshold Schemes with Minimum Pixel Expansion

References

1. M. Naor and A. Shamir, "Visual Cryptography," Lecture Notes in Computer Science, Vol. 950, 1994, pp. 1–12.
2. P.-A. Eisen and D.-R. Stinson, "Threshold Visual Cryptography Schemes with Specified Levels of Reconstructed Pixels," Design,



Visual Cryptography. Figure 1. Visual cryptography concept demonstrated using a (2, 2)-threshold scheme.



Visual Cryptography. Figure 2. A (2,2)-VSS scheme constructed using 2×4 basis matrices with $m_1 = 2$ and $m_2 = 2$: (a) secret input image, (b, c) binary shares, (d) output image visually decrypted by stacking the share images shown in (b, c).