Volume 1, No. 2, July-August 2010

ISSN No. 0976-5697

International Journal of Advanced Research in Computer Science

RESEARCH PAPER

Available Online at www.ijarcs.info

An Efficient Method to find Video Objects

Dr. M. Mohamed Sathik
Reader in Computer Science,
Sadakathullah Appa College,
Thirunelveli, India
mmdsadiq@gmail.com

Ms. M. Parveen
Research Scholar in Computer Science,
Bharathiyar University,
Coimbatore, India
parveenjagir@gmail.com

Ms. P. Peer Fatima*
Research Scholar in Computer Science,
M.S.University,
Thirunelveli, India
pfatimabasha@gmail.com

Abstract: Modern Computer Technology, together with the proliferation broadcast channels and of video-based surveillance systems, has enabled us to produce vast amounts of both raw and processed video data. The potential uses of this data are many and varied. Monitoring and mining of the contents of this is already huge, rapidly growing mass of data calls for the development of major computational resources and the development of sophisticated video understanding techniques. Video extraction / video mining is a highly challenging tasks. It deals with the identifying / capturing of object of interest from the video. Some common area where this video mining is useful are: identifying the pitch of ball inside / outside the boundary, captures sachin's role in the entire match, analyses the winner in the running race, etc. In this article, we propose a suitable technique for capturing the moving object in video using frames. We capture the object and place it in frame in a sequence manner. The background in the object is eliminated by some technique. Thus the foreground image is well displayed. After elimination of the background, the objects are compared with each other. Finally, the difference in the frames is captured. Thus the particular object is captured from the video. The objective is to monitor and capture the object in a random manner. The framework for identifying the object is by recording through a stationary camera, and captures the sequences of objects. Dynamic objects can be captured by both background elimination and background registration techniques. Thus we propose the suitable methods to mine the object from the video in a more easy way. The background registration method uses background subtraction which improves the adaptive background mixture model and makes the system learn faster and more accurately, as well as adapt effectively to changing environments.

Keywords: Background elimination, background registration techniques, modern computer techniques, monitoring and mining of the content, Object identification, stationary background, stationary camera.

I. INTRODUCTION

Now a day, the object-based MPEG4 standard is becoming growingly attractive to various applications in area such as the Internet, video editing, and wireless communication because of its object-based nature. For instance, in video editing, it is the object of interest, not the whole video, which needs to be processed; in video transmission, if the bandwidth of the channel is limited, only the objects not the background are transmitted in real time. Generally speaking object-based video processing can simplify video editing, reduce bit rate in transmission, and make video search efficient.

The purpose of this workshop will be to survey available and potential technologies for video monitoring and mining (and in general methods of fast and efficient content-based analysis of video streams) and to identify promising directions for research in this challenging area. Specific topics to be covered will be analysis of camera motion and scene activity; temporal segmentation; content-based classification, indexing, and retrieval; representation, browsing, and visualization.

We will also investigate related mining problems having to do with audio mining, seismic data mining, and cross-modality mining. Audio mining applications of interest include speech recognition; seismic data mining applications include identification of potential new sources of oil and gas and detection of earthquakes and/or nuclear tests. Cross-modality issues arise for example in problems involving identification from both video and speech.

Particle filter is an important technique in the video mining concept. This type of filtering is called OBJECT EXTRACTION or OBJECT MINING. The object may be any person or thing. The objective of this article is to play the video and then extract the particular object and capture it by selecting the object. This extraction may be done for single object or for multiple objects. The main scope of this article is to propose a secure, robust and object-based video authentication solution.

In previous days, the object can be viewed through the camera. When the number of camera increases, it requires more operators to trap the object. So it is a tedious process to do this job with the camera. It is possible to analyze and propose a suitable method to extract a lot of information in an easy way. One such technique is called **VIDEO MINING.**

Video mining can be defined as the unsupervised discovery of patterns in audio visual content. Due to the success of data mining techniques, video mining techniques will be emerged. Automated surveillance systems typically use stationary sensors to monitor an environment of interest. This detects the object in the video frame. The objective of

the video mining is to extract the interesting objects from the video based on its re-occurrence.
Video mining is done by using stationary camera to view the object. But it is a difficult process to capture the object and it is inexpensive too. Manually, reviewing the large amount of data they generate is often impractical.

There are some pervasive challenges here. A huge amount of data must be efficiently stored. Once stored, there must be searched for the efficient retrieval algorithms. It must also be possible to retrieve different kinds of data stored in possibly different formats. Some of the applications will require real-time response and action based on an incoming data stream. Some of the applications The applications and potential applications of video mining include:

1. Monitoring of (possibly remote) surveillance cameras in theft protection, fire protection, care of bedridden patients and young children
2. Automatic recognition of suspicious people in large crowds
3. Automatic checking to identify passengers entering an airplane, bus, or public building (to verify entry authorization)
4. Automatic quality control in manufacturing processes
5. Robotic vision
6. Retrieval of archived video clips to illustrate a news breaking story
7. Retrieval of suspicious activities in prerecorded video surveillance sequences
8. Browsing of DVD and set-top box recordings intelligent fast-forward techniques
9. Detection of multi-lingual scene text to determine origin of broadcast
10. Video search engines for web browsers
11. Ranking of video clips by relevance in web query results classification of videos into genres for search pruning
12. Detection of crowd patterns for mob control
13. Detection of traffic patterns for traffic understanding management

We also propose a method to mine the video in a better way and to extract more useful information from the video.

The remainder of this workshop is organized as follows: Section-II states about the related work regarding the topic. Section-III describes about the methodology with the algorithm and the section-IV describes the experimental results of our work. Finally, in Section-V we conclude this article

II. RELATED WORK

A brief survey of the related work in the area of the video mining is presented in this section.

The adaptive observation model arises from the adaptive appearance model as inspired by [3] and also the appearance model is based on phase information derived from the image intensity whose computation is quite time-consuming. The direct embedding of such model in a particle filters further increase the computational burden. Thus, they use the intensity based appearance model.

The occlusion is usually handled in 2 manners.

1. One way is to use joint probabilistic data associative filter (JPDAF) as in [4];and
2. The other is to use robust statistics as in [5]. The face detector is used for the face sequence in [6].

The content-based learning content management system (CBLCMS) is carried out in [7]. Through this system, video content can be segmented by an automatic shot and scene retrieved technology and stored in the Learning Asset Database (LADB). The concept of metadata information retrieved from multimedia retrieval system embedded in the framework model of the E-Learning content object, and expect that these data-oriented metadata can be used and read in SCORM standard learning management system is proposed. In this research, related scene cutting as well as regarding the news abstract in the television news retrieval system the part.

In the article [8], it addresses the problem of extracting video objects from MPEG compressed video, using motion vectors. Also the method for automatically estimating the number of objects and extracting independently moving video objects using motion vectors. Though much work is being done in the area of motion based video object segmentation in pixel domain [15], [16], [17], [18] very little work has been carried out in the area of compressed domain VOP extraction.

The work for automatic generation of textual metadata based on visual content analysis of video news using the methods for semantic object detection and recognition, concerned with two semantic objects: faces and TV logos as in[9].

Sikora T, in [25], used the concept for intelligent signal process and content-based video coding. Here an image scene consists of video objects and the attempt is to encode the sequence that allows separate decoding and construction of objects.

Nack et al [26] and Salembier et al [27] have discussed Multimedia content description related to the generation of region based representation with respect to MPEG-4 and MPEG-7.

Now, we consolidated all the above papers and create it with more qualitative and in technical wise using more features.

III. METHODOLOGY

A. Proposed Method

The aim of the proposed method is to develop an application to extract the object from the video. In this case it is possible to use the method to extract the object in a suitable manner.

The summary of this method is as follows: The video content is observed and then it is transformed into the frames such that one object is captured per frame. The object can be inserted into the frames only after by segmenting the frame into many sections and then arrange it in sequences. So that, when the video file is transformed into the frames, the object situated in a sequence manner. This process is very useful for us to compare the frame and the object. The video mining is done by taking the following cases into observation:

1. Transformation into Frames
2. Frame Difference
3. Eliminating the stationary background
4. Comparing the frame object
5. Object Detection
6. Noise Reduction
7. Object Identification
8. Resultant video mining

B. Description

(i) Transformation into Frames

The frame is splited into several segments and then transforms the video into the frame such that, each frame contains an object. Consider the input video clip as N and each of the frame objects as F1, F2,……Fn, such that

$N = \sum_{i=1}^{n}(Fi)$

(ii) Frame Difference

The frame difference is observed. This can be done by comparing the consecutive frames and analyze the difference between them.

(iii) Eliminating the stationary background

After finding the frame differences, the next step is to removal of background from the object in each frame. Since each background in the object will make the object looks differently, it may lead to a wrong evaluation. So it is necessary to eliminate the background of the object. Thus the frame contains only the object.

(iv) Comparing the Frame Object

The next step is to compare each frame and extract the object. This can be done by considering some techniques such as pixel comparison / image comparison.

(v) Object Detection

After eliminating the background from the object, the next step is to detect and trap the object. The elimination of background from the object helps to retrieve the dynamic foreground object. Tracking the object through the video sequence is by comparing each frame with one another.

(vi) Noise Reduction

The processing technique is applied in the video mining to reduce and eliminate the unwanted noise in the background and to trap clear results.

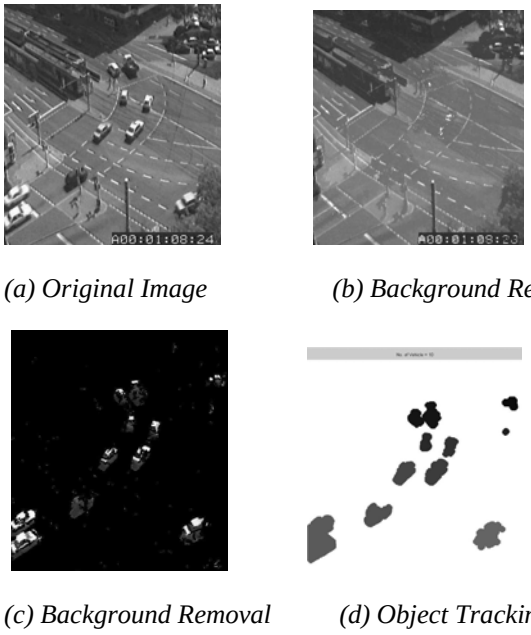
(vii) Object Identification

After reduce the noise levels and eliminating the background, now the object is identified easily by comparing the pixel position of the object in the frame. If the pixel position of the object is greater than a certain threshold, then those pixels are replaced by the pixels of the original frame.

(viii)Resultant Video Mining

After comparing the pixel position of the image, the object is identified in the frame. Now each frame contains single object within it. The user selects a particular object from the video to analyze the role of that object in the entire video. This comparison is made by selecting the frame in a random manner either by selecting all the even frames or by selecting only the odd frames.

After selecting the sequence of frames, the object is compared and extracted. Due to this action, the object is captured successfully in less time. The rest of the object is stored separately. The resultant object is then converted into Gray Scale object. But this process does not produce actual objects. The reverse process of converting the Gray Scale object into black and white indirectly identifies the object and thus the object is extracted from the video.



(a) Original Image *(b) Background Registration*

(c) Background Removal *(d) Object Tracking*

Figure 1: Video Mining

The proposed method consists of an algorithm. All the required validation processes will be taken in consideration by the proposed method. The algorithm is as follows:

C. Algorithm

Read the video
Split N frames with regular intervals
Convert the video file and place it in N frames with one object per frame
Analyze the consecutive frames to identify the frame differences
Eliminate the background from the frame
Extract the object from the frame
Comparing extracted object and detect the object
Reduce the noise level in all N frames
Identifies the object by comparing the pixel position of the object
Get the user input for mining the particular object from the video
Convert the RGB image into Gray Scale object
Convert the Gray Scale image into Black and White object
Compare each object with the user input
If both objects coincides then
    Extract the object from the video
End if
Video mining is done successfully.

D. Algorithm Explanation

Video mining is the process of extracting the object from the video. First of all, the video is transformed into the frames. Then find the differences in the frames by analyzing consecutive frames. Then eliminate the background of the frame and view the dynamic object in a clear way. And then, reduce the noise level in the object.

The object is compared with each other by matching the pixel position of the object. Then convert the RGB image into Gray Scale image and then convert into Black and White image.

Then compare the object with the user input and extract that object from the whole video. Thus the object is ex-

tracted from the video. Thus video mining is done successfully.

E. Comparison Chart

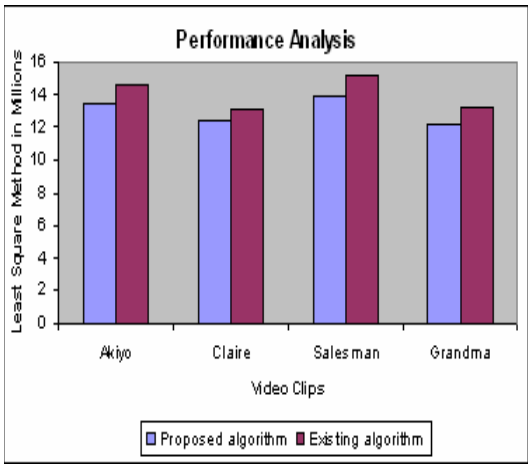


Figure 2: Comparison Chart between Existing algorithm and the Proposed Algorithm

IV. EXPERIMENTAL RESULTS

The implementation of the Video mining technology is an interesting one. This is very useful in many real-time environments. When the video plays, we can trap the particular object and made several differentiation among those objects. Some of the examples where this video mining is implemented are as follows:

1. In the field of cricket, it is easy to trap the role of Sachin Tendulkar in the whole match. This is possible, by selecting the Tendulkar image in the video. This will identify all the locations where Tendulkar been placed.
2. In running race, car race or any other races, it is easy to detect the first position when using video mining technique. By using this technique, one can capture the winner accurately by viewing the first one who crosses the final point.
3. This technique is also useful to control traffic, by tracing the person or vehicle those/which will cross the white line when the signal is red.
4. Through this Object Tracking or Object Mining, it is the easy way to deduce whether the batsman got six or four by tracing the ball crossing the boundary.
5. It is also possible to detect whether the ball goes for catch after pitching.

We implement the video mining technique only for the single object. By tracing the single object, the video mining extracts all the areas in which the selected object is identified. It is more useful in gaining information about the particular object in the entire video. One of the examples is to trace the Sachin's role in the Indo-Pakistan Match.

We also present a method for simultaneous tracking and recognition of visual objects from video using a time series model with stochastic diffusion. Specifically, by modeling the dynamics with a particle filter we are able to achieve a much stabilized tracker and an accurate recognizer when confronted by pose, scale and illumination variations. We have tested our tracker on real-time face detection (invariant to scale/rotation) as well as tracking vehicles from ground-level views as well as aerial surveillance video of tanks seen from oblique (affine) views. This is by using multiple objects. In the following video mining, at a time, the user can gain information on various objects.

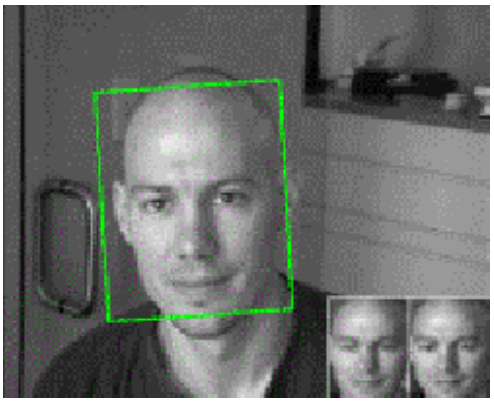


Figure 3: Single Object Tracking

Figure. 4: Multiple Object Tracking

V. CONCLUSION

Video mining techniques is successfully implemented both for single object and for multiple objects. Using the video mining technique, the area of research is interesting and user-friendly.

In image retrieval systems for specialized domains, image segmentation can focus the search very precisely. Retrieval in broad, PACS like databases needs divergent algorithms to extract the most important parts of the image for indexing and retrieval. We present such an algorithm that works completely automatic and quickly and as a consequence can be applied to very large databases. Some of the recognized problems might be particular for our setting but they will appear in a similar fashion in other teaching _les. The particular problems will need to be detected for any other collection.

Our solution is optimized to have very few images where too much is cut as this could prevent images from being retrieved. This algorithm can be implemented in a massively parallel environment, which gives the hope for online object segmentation. The above segmentation is performed only by taking the sparse motion vectors as input. The segmentation can be further improved by considering the spatial intensity values of the image frame, which can be obtained by partially decoding the MPEG stream.

We have improved our simultaneous tracking and recognition approach proposed. More complex models, namely adaptive appearance model, adaptive-velocity transition model, and intra and extra-personal spaces model, are introduced to handle appearance changes between frames and between frames and gallery images. The fact that gallery images are in frontal view is enforced too. Experimental results demonstrate that tracking is stable and the recognition performance has improved.

VI. REFERENCES

[1] Sikora, T., the MPEG-4 Video Standard Verification Model, IEEE Transactions, Circuits Systems, Video Technology, vol. 7, pp. 19-31, Feb. 1997.

[2] Nack F. and Lindsay A. T., Everything you wanted to know about MPEG-7: Part 2, IEEE Multimedia, vol. 6, pp. 64-73, Dec. 1999.

[3] A.D.Jepson, D.J.Fleet, and T.El-Maraghi, "Robust Online Appearance model for visual tracking. Proc. Of CVPR, 1:415-422,2001

[4] C. Rasmussen and G.D.Hagar, "Probabilistic data association methods for tracking complex visual objects" IEEE Transactions on Pattern Analysis and Machine Intelligence, 23(6):560-576,2001

[5] P.J.Habel, Robust Statistics, Wiley, 1981.

[6] P. Voila and H. Jones, "Robust Real-Time Object Detection Second Intl. Workshop on stat and comp. Theories of vision 2001"

[7] "Video Learning Object Extraction and standardized metadata"- Tzong. Per wu, ying-yuan yeh, yu-Hingchou, Chinese Culture University, Dept. of information Communications, National Taipei College of Nursing, Dept. of Nursing, wuder@faculty.pccu.edu.tw, corn.dennis@gmail.com

[8] "Compressed Domain motion segmentation for video object extraction"- R.Venkatesh Babu and K.R. Ramakrishnan.

[9] "Automatic textual annotation of video news based on semantic visual object extraction"-Nozha Baljemaa, Francois Fleuret, Valerie Ciouet, Hichem Sahbi INDRI-IMEDIA Research Group

[10] Salembier P. and Marques F., Region-based Representations of Image and Video: Segmentation Tools for Multimedia Services, IEEE Transactions, Circuits Systems, Video Technology, vol. 9, pp. 1147-1169, Dec. 1999.

[11] Visual Tracking and Recognition Using Appearance-Adaptive Models in Particle Filters-Shaohua Zhou, Rama Chellappa, Baback Moghaddam

[12] Automatic Textual Annotation of Video News Based on Semantic Visual Object Extraction- Nozha Boujemaa, Francois Fleuret, Valerie Gouet, Hichem Sahbi, INRIA - IMEDIA Research Group - Domaine de Voluceau - BP 105 - F78153 Le Chesnay Cedex

[13] Automated Object Extraction for Medical Image Retrieval Using the Insight Toolkit (ITK)- Henning Muller, Joris Heuberger, Adrien Depeursinge, Antonie Geissbuhler University and Hospital of Geneva, Medical Informatics, Geneva, Switzerland, henning.mueller@sim.hcuge.ch

[14] A Secure and Robust Object-Based Video Authentication System- Dajun He, Qibin Sun, Qi Tian, Institute for Infocomm Research (I2R), 21 Heng Mui Keng Terrace, Singapore 119613, Email: tian@i2r.a-star.edu.sg

[15] Noel Brady and Noel O'Connor, "Object detection and tracking using an EM-based motion estimation and segmentation framework," in Proceeding of the IEEE International Conference on Image Processing, 1996, pp. 925-928.

[16]David P. Elias, The motion Based Segmentation of Image Sequences, Ph.D. thesis, Trinity College, Department of Engineering, University of Cambridge, Aug. 1998.

[17] Nuno Vasconcelos and A. Lippman, "Empirical Bayesian EM-based motion segmentation," in Proceedings of the IEEE

[18] Philip H. S. Ton; Richard Szeliski, and P. Anandan, "An integrated bayesian approach to layer extraction from image sequences," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 23, no. 3, pp. 297-303, Mar. 2001.

[19] Wang D., Unsupervised Video Segmentation Based onWatersheds and Temporal Tracking, IEEE Transactions, Circuits Systems, Video Technology, vol. 8, pp. 539-546, Sept. 1998.

[20] Compressed Domain Motion Segmentation for Video Object Extraction, R. Venkatesh Babu and K. R. Ramakrishnan, Department of Electrical Engineering, Indian Institute of Science, Bangalore 560 01 2, India.

[21] Automated Object Extraction for Medical Image Retrieval Using the Insight Toolkit (ITK), Henning M□uller,Joris Heuberger, Adrien Depeursinge, Antoine Geissb□uhler, University and Hospitals of Geneva, Medical Informatics, Geneva, Switzerland

[22] Video Learning Object Extraction and Standardized Metadata Tzong-Der Wu Ying-Yuan Yeh Yu-Ming Chou Chinese Culture University Department of Information Communications National Taipei College of Nursing Department of Nursing

[23] Automatic Textual Annotation Of Video News Based on Semantic Visual Object Extraction Nozha Boujemaa, Francois Fleuret, Valerie Gouet, Hichem Sahbi INRIA - IMEDIA Research Group - Domaine de Voluceau - BP 105 - F78153 Le Chesnay Cedex

[24] Adaptive Visual Tracking and Recognition using Particle Filters Shaouha Zhou _ Rama Chellappa _ Baback Moghaddam y TR2003-95 July 2003

[25] Sikora, T., The MPEG-4 Video Standard Verification Model, IEEE Transactions, Circuits Systems, Video Technology, vol. 7, pp. 19-31, Feb.1997.

[26] Nack F. and Lindsay A.T., Everything you wanted to know about MPEG-7: Part 2, IEEE Multimedia, vol. 6, pp. 64-73, Dec. 1999.

[27] Salembier P. and Marques F., Region-based Representation of Image and Video: Segmentation Tools for Multimedia Services, IEEE Transactions, Circuits Systems, Video Technology, vol. 9, pp. 1147-1169, Dec.1999.

[28] Flickner, M., Sawhney, H., Niblack, W., Ashley, J., Huang, Q., Dom, B., Gorkani, M., Hafner, J., Lee, D., Petkovic, D., Steele, D., Yanker, P.: Query by Image and Video Content: The QBIC system. IEEE Computer 28 (1995) 23{32

[29] Bach, J.R., Fuller, C., Gupta, A., Hampapur, A., Horowitz, B., Humphrey, R., Jain, R., Shu, C.F.: The Virage image search engine: An open framework for image management. In Sethi, I.K., Jain, R.C., eds.: Storage & Retrieval for Image and Video Databases IV. Volume 2670 of IS&T/SPIE Proceedings., San Jose, CA, USA (1996) 76{87

[30] Graham, M.E., Eakins, J.P.: Artisan: A prototype retrieval system for trademark images. Vine 107 (1998) 73{80

[31] M□uller, H., Rosset, A., Vall_ee, J.P., Geissbuhler, A.: Integrating content{based visual access methods into a medical case database. In: Proceedings of the Medical Informatics Europe Conference (MIE 2003), St. Malo, France (2003)

[32] Tagare, H.D., Ja_e, C., Duncan, J.: Medical image databases: A content{based retrieval approach. Journal of the American Medical Informatics Association 4 (1997) 184{198

[33] B. Schneier, Applied Cryptography: Protocols, Algorithms, and Source Code in C, Wiley, New York, NY, USA, 2nd edition, 1996.
[34] F. Bartolini, A. Tefas, M. Barni, and I. Pitas, "Image authentication techniques for surveillance applications," Proceedings of the IEEE, vol. 89, no. 10, pp. 1403–1418, 2001.
[35] A. Piva, R. Caldelli, and A. De Rosa, "A DWT-based object watermarking system for MPEG-4 video streams," in Proc. IEEE International Conference on Image Processing (ICIP '00), vol. 3, pp. 5–8, Vancouver, BC, Canada, September 2000.
[36] C.-Y. Lin and S.-F. Chang, "Robust image authentication method surviving JPEG lossy compression," in Storage and Retrieval for Image and Video Databases VI, I. K. Sethi and R. C. Jain, Eds., vol. 3312 of Proc. SPIE, pp. 296–307, San Jose, Calif, USA, December 1997.
[37] J. Dittmann, A. Steinmetz, and R. Steinmetz, "Content-based digital signature for motion pictures authentication and content-fragile watermarking," in Proc. IEEE International Conference on Multimedia Computing and Systems, vol. 2, pp. 209–213, Florence, Italy, July 1999.
[38] M. P. Queluz, "Towards robust, content based techniques for image authentication," in Proc. IEEE 2nd Workshop on Multimedia Signal Processing, pp. 297–302, Redondo Beach, Calif, USA, December 1998.
[39] N. V. Boulgouris, F. D. Koravos, and M. G. Strintzis, "Selfsynchronizing watermark detection for MPEG-4 objects," in Proc. IEEE 8th Conference on Electronics, Circuits and Systems (ICECS '01), vol. 3, pp. 1371–1374,Masida,Malta, September 2001.
[40] C.-S. Lu and H.-Y. M. Liao, "Video object-based watermarking: a rotation and flipping resilient scheme," in Proc. IEEE International Conference on Image Processing (ICIP '01), vol. 2, pp. 483–486, Thessaloniki, Greece, October 2001.

.