Biometric Access Control for Digital Media Streams in Home Networks

Peter Corcoran, *Member, IEEE,* Claudia Iancu, Frank Callaly and Alex Cucos

***Abstract* —** *Methods for controlling access to digital media content in the home environment using biometric sensing techniques are described. A wireless rebroadcast system for IP-TV content was retrofitted with a combination of fingerprint authentication and face recognition modules. Family members are enrolled and new individuals may be added to the system by a master user. Access to content requires fingerprint based authentication from a wireless PDA and the system also checks dynamically if it can recognize enrolled uses amongst those "sitting on the couch".*

***Keywords* — Biometrics, Multimedia, Digital Media Encoding, Home Networks.**

I. Introduction

The dissemination of such new technologies which facilitate the copying, compression and sharing of digital content over network connections has created problems for both the music industry and Hollywood in recent years. As a society, if we value these services there is clearly a pressing need to manage and account for the copying and redistribution of digital multimedia.

In recent times the Recording Industry Association of America (RIAA) has taken the somewhat heavy-handed approach of initiating broadly targeted legal actions against individuals who are involved in the sharing of digital content. Despite such actions there is evidence that allowing controlled copying and sharing of digital content can lead to market growth and improved sales. Thus a key part of the challenge for content providers in today's digital age is to provide mechanisms which allow copying and rebroadcast of digital content within the home environment combined with limited sharing of digital content to friends and family members but which restrict commercial piracy.

An initial approach was presented in [1] where the authors describe a method of encrypting digital content with a public key generated from biometric data associated with the owner of the content. This work addressed the problem of allowing consumers "fair use" rights while at the same time restricting the illegal piracy of digital media.

Biometric techniques for identifying and authenticating people are well known. Techniques include face recognition [2], fingerprint [3] and voice authentication [4]. There are three main steps in the biometric evaluation process: firstly the biometric data must be obtained using a sensing apparatus; secondly the data must be analyzed and thirdly compared with a database of signatures.

In this paper we focus on applying the techniques of [1] in a practical home environment. The idea was to build a working home network with biometric control components which require simple authentication in order to initiate access to content and which continually monitor the people viewing content. Thus we combine an initial authentication event with ongoing background monitoring to confirm that at least one enrolled party continues to view the content. The goal was to explore new modes of content access management which can be achieved over a home network by employing smart peripherals.

II. Media Stream Rebroadcasting System

From earlier research work [5] we have developed a rebroadcast appliance which can obtain media streams from a number of conventional sources. The key subsystems of this appliance and its relationship to a local 802.11 home network are shown in ***Fig 1***. This standard configuration allows conventional TV signals to be captured and digitized in real-time as an MPEG2 stream. Streams can also be obtained from an IPTV source which employs multicast group techniques over a wide area network, or from a standard DVD player. Transcoding to MPEG4 or DivX can also be realized, although this is not currently feasible in real time.

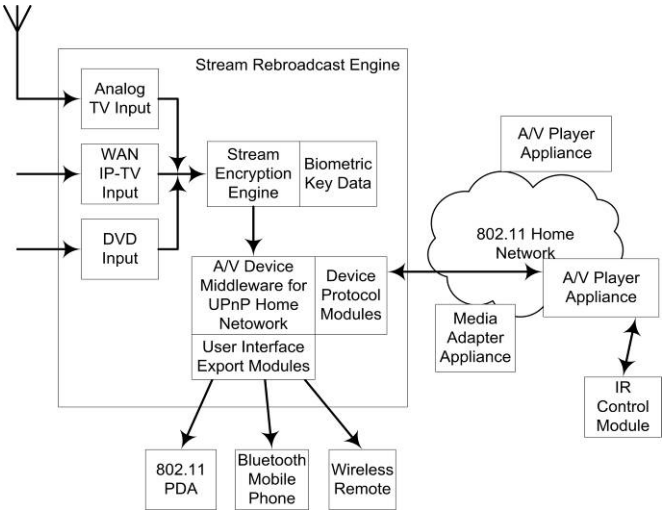

Fig 1: Content Rebroadcasting using Metadevice Middleware

An overview of the biometric stream encoding techniques employed in this paper was given in an earlier published work [6]. The principle goal of the present research is to

This work is supported under the CFTD Program (CFTD/2004/206) of Enterprise Ireland under the National Development Plan.

Peter Corcoran is with the Dept. Electronic Engineering, National University of Ireland, Galway (e-mail: peter.corcoran@nuigalway.ie).

Claudia Iancu, Frank Callaly and Alex Cucos are with the Dept. Electronic Engineering, Natuonal University of Ireland, Galway (e-mail: claudia, frankc and cucos@wuzwuz.nuigalway.ie).

Manuscript received July 2, 2007

investigate practical issues surrounding the implementation of a working system which uses biometric inputs to control user access to content. In particular we are interested in making the workflow model of content access as transparent as possible with an end goal that users should be unaware when key control actions require biometric confirmation.

III. Review of Authentication Techniques Evaluated

In the system described in this paper we have used two principle modes of authentication: (i) fingerprint based analysis is employed where active user input is required and (ii) face recognition of users, employed where passive authentication is sufficient. The use of face recognition as a primary mode of authentication was also investigated, but it is less reliable than fingerprint based analysis.

Before using the system a user must enroll their fingerprint and face data with the system. It is worth commenting that the weakness with any means of biometric authentication for consumer applications lies in the enrollment process. If an accurate enrollment is not completed then the system will perform unreliably and users will perceive it as a burden.

A. Fingerprint Matching Techniques

The fingerprint authentication subsystem is described in detail in a companion paper [32]. Here we will only focus on issues relating to matching 2D *minutiae* (or key-point) patterns between a newly acquired fingerprint and a set of successfully enrolled fingerprints. The problem of fingerprint matching thus reduces to finding an algorithm which can quickly determine if a suitable match exists between the newly determined 2D point pattern and a set of previously enrolled patterns. A number of different approaches exist:

1) Affine Pattern Matching Techniques

Quite a few different algorithms have evolved, particularly from the field of *image registration* to determine an optimal match between two 2D point patterns. Amongst the most recent approaches we mention the *iterative closest point algorithm* (ICP) described in [8]. As with many of the following variants the ICP algorithm assumes a rigid Euclidean transformation between corresponding point sets.

The basic ICP algorithm has been extended by a number of authors. For example, in [9] it is extended to handle correspondence between a point and a tangent plane to overcome the lack of an exact correspondence between the two point sets; in [11, 13] the algorithm is made more robust to the influence of outliers and features lacking correspondence; in [10] a weighted least mean error estimate is employed and in [12] a metric is employed which trades off distance against feature similarity based on local shape invariance. For real-time applications [14] examines performance improvements to the basic algorithm.

Another recent approach to such problems is that of *thin plate spline* (TPS) matching algorithms [15, 16] which are robust to false positive candidate spots and non-affine warpings between the candidate point sets. This algorithm works by iteratively performing a joint estimation of the correspondence between two sets of points which are to be registered, and the TPS mapping between the images that the points lie in.

Although we have tested some of the above algorithms and developed our own variant which is modified to accommodate aspects specific to fingerprint *minutiae* point sets we reached a conclusion that such algorithms are flawed for our purpose as they assume a correspondence between the two point sets to be matched. Thus they are not optimized to indicate quickly when point sets are unlikely to allow a good match.

2) Ridge Field Pattern Matching Approaches

An alternative approach to fingerprint pattern matching is provided by analyzing the *ridge-field pattern* of an acquired fingerprint. This provides useful global information which is additional to the *minutiae* points and which is relatively robust to the quality of image acquisition [22]. When combined with local *minutiae* based features the use of this orientation field data can provide both accurate and robust matching [21]. However the algorithms proposed by these authors are too slow for practical use in embedded CE systems.

3) Minutiae Triangle Descriptors

A number of authors have turned their attention to an alternative approach to the matching of fingerprint patterns [17-20] based on the use of local triangular features formed from nearest-neighbor *minutiae* groupings. In [17] the author builds a connected graph composed of triangles formed form nearest-neighbor *minutiae* with shared edges. This graph is extended outwards from an initial matched set of minutiae. By allowing for small localized variations in the lengths of the shared edges this matching technique can account for global variations of us to 45% across the entire fingerprint. The authors of [18] use localized matching of *minutiae pairs* combined with a global matching. They use the best matched local structure pair to determine a global translation and rotation for the remaining *minutiae* and perform their overall decision based on the resulting match of the global *minutiae* set. In [19] the authors employ fuzzy similarity measures for determining matches between individual local triangle features and combine these to obtain a global measure of similarity between two fingerprints. Finally, the authors of [20] build a set of triangle descriptor feature vectors for each fingerprint and then proceed to determine if the triangular frameworks for each fingerprint correspond closely enough to declare a match. Their technique is differentiated from that of [18] as it is not necessary to perform a global alignment of the *minutiae* patterns.

These approaches [17-19] are broadly equivalent to the affine point-matching problem when the global triangular structure is analyzed, although the local triangular features can provide useful additional data to improve the speed and accuracy of the global pattern matching process. However

we note that where additional local *ridge-field* data is used the discriminating power of these local triangular features is somewhat improved and, more importantly, these features are essentially rotation-invariant a described in [20]. An example of this modified triangular descriptor is given in ***Fig 2*** below.

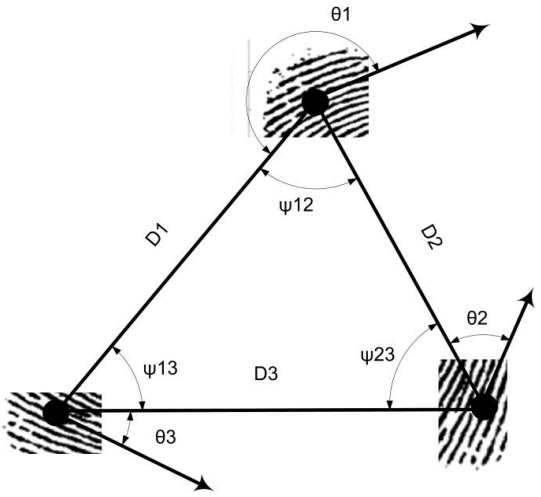

Fig 2: Local Triangle Descriptor formed by 3 Minutiae Points

One weakness in the use of such triangular descriptors, and indeed for all of the above techniques is that they are susceptible to erroneous acquisition data. If a single *minutiae* point is omitted, or is incorrectly detected as two adjacent *minutiae* then the pattern matching algorithm will often fail. This is not a problem for most authentication scenarios as the user can simply re-apply their fingerprint, paying more attention to the pressure and duration of the acquisition process. However in a consumer application it is desirable that even if the acquisition process is poorly executed there is still a high probability of a successful outcome.

For example, if content recording or playback is conditional on obtaining a correct response a user would become frustrated if they must press the "play" button 3 or 4 times to actuate playback of a movie.

B. Face Recognition Techniques

Face recognition can be employed in two principle modes for our IPTV system: (i) it can serve simply as a means of periodically authenticating that the people viewing some content are members of a "viewing group"; (ii) it can form part of the core authentication process either in place of, or as a compliment to fingerprint based authentication.

For the first usage mode we can actually manage without a permanent face recognition database if we accept that when a movie starts playing the original set of viewers are genuine and are verified by the person providing fingerprint authentication process that initiates movie playback. In the case of the second usage mode then we need to consider a face enrollment process for users. The choice of usage mode will also impact on the face recognition techniques that are most useful and reliable.

There are many known approaches to face recognition developed in the literature, most of them with very good results on certain databases or when employed in a particular context or scenario. However there is no global solution to the face recognition problem. Each of the known algorithms has particular advantages and disadvantages.

Our application context is further complicated because we would like the authentication process to be somewhat transparent to the user: we need to acquire facial images for authentication in variable poses and illuminations; the algorithm has to be simple enough to implement on consumer devices yet robust enough when working with image set where large variations are likely to appear. Because we want to apply the face recognition in a passive mode we cannot specify constrains regarding how the face images should be acquired. Some of these considerations were the subject of earlier work [2, 23, 26-28] and the interested reader is referred to these, particularly [28], for further details.

1) DCT Based Face Recognition

One of the simplest and fastest approaches for face recognition is to use as similarity measure between two faces, the distance between their DCT spectrum, or a part of the spectrum. However this approach has proven to be quite sensitive when applied on facial regions with significant variations in illumination and pose. Nevertheless, when some preprocessing is employed to minimize such variations [28] it is one of the more useful forms of face recognition for distinguishing between the end users for consumer applications due to its relative simplicity and simplified training requirements.

2) PCA Based Face Recognition

Another efficient method for face recognition, *principal component analysis* (PCA), presents the same disadvantage when applied on common user databases. The PCA is based on projecting the faces using a set of basis vectors corresponding to the largest variations inside the face collection and computing the distances in this feature space as similarity measure. PCA algorithm is more complex than DCT but it has been proven more efficient for face recognition by many researchers. A particular disadvantage of the PCA approach is that is data dependent, which means that every time the face collection changes the PCA analysis has to be repeated to re-compute the main basis vectors. This retraining is an important limitation of PCA for our application.

3) Image Preprocessing for PCA/DCT Recognition

As stated above these two techniques are quite sensible to variations especially variations in pose and illumination. One solution to increase the robustness of the recognition is to minimize/normalize these variations.

For illumination variations, an easy technique commonly used is Histogram Equalization performed on the face regions. HE performs well for relatively uniform and small variations in illumination. For larger variations other algorithms have to be considered, like CLAHE (contrast limited adaptive histogram equalization) [24].

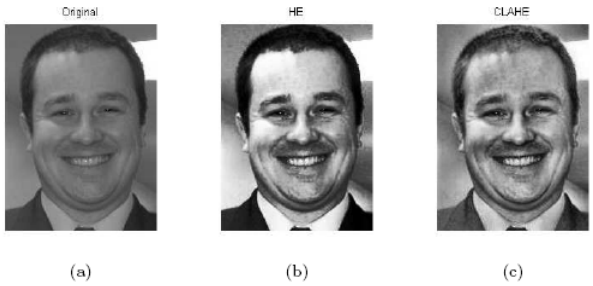

Fig 3: Example of using CLAHE image normalization: (a) original; (b) using histogram equalization and (c) CLAHE.

Another important factor that influences the recognition accuracy is face orientation. In order to normalize these types of variations face modeling techniques have to be considered. One such technique successfully tested on the combined PCA/DCT recognition approach is AAM [25]. Unfortunately the complexity of this modeling technique does not recommend it for embedded implementations just yet.

4) HMM Face Recognition

Another main disadvantage for both PCA and DCT is that even if the system is trained with several pictures of the same face in different conditions, a robust model cannot be built. A good solution for this problem is the use of HMMs as the basis for a face recognition algorithm. Even though HMMs are more complex to implement, the models derived during the initial training process are more robust so the results are much better in real world conditions.

Hidden Markov Models are a set of statistical models used to characterize the statistical properties of a signal. HMMs consist of two interrelated processes: an underlying Markov chain with a finite number of states, a state transition probability matrix and an initial state probability distribution, and a set of observations, defined by the observation density functions associated with each state. The HMMs classification is based on modeling the pattern to be recognized as a process represented by a succession of states. The states represent parts of the pattern and by analyzing them we can obtain sets of observations which are used to train models and are subsequently used for classifying newly acquired face images.

The face recognition system we used is based on a variant of 2D HMM called embedded HMM. The embedded HMM is a generalization of the classic HMM, where each state in the one dimensional HMM is itself an HMM. Thus, an embedded HMM consists of a set of super states along with a set of embedded states. The super states model the two dimensional data along one dimension, while the embedded HMM models the data along the other dimension.

Considering the fact that the facial features in an image occur in a natural order when the image is scanned the same way we can assign the same states in the model for the same regions in the image. One advantage of using HMMs for face recognition is their robustness to small rotations and small variations in face orientation. In order to decrease the effects of variations in illuminations, we applied CLAHE algorithm on each image before training/testing phase.

5) Comparative Test Analysis

We performed a series of tests to analyze the performances of the face recognition algorithms. For that we used a database which contains 560 pictures corresponding to 56 persons, with high variations in pose, expression and illumination. Initially we tested the DCT approach and extension of the classical PCA approach on wavelet decomposition of the original image. For that we resized all images to 32x32 pixels and transformed them to gray scale images. To limit the influence of the illumination variation we applied CLAHE on the cropped face regions. The recognition accuracy was between 47.86% and 87.14%, depending on how many images per person were used in the training stage. We also used a combination approach between the two algorithms and the recognition rate increased to 88.21%.

Next we tested the HMMs face recognition approach. Prior to analysis each image was resized to 128x128 pixels, this increasing the recognition rate with almost 5%. We used 1 to 5 images for training and the rest of 5 images for testing. The results of HMM algorithm with DCT features as observation vectors and CLAHE for illumination normalization are shown in table 1. It can be noted that a good accuracy can be achieved when only 2/3 images per person are used to build the models.

No training faces/ No testing faces	1/5	2/5	3/5	4/5	5/5
Recognition rate (%)	68.57	78.2	86.1	92.9	95.7

Table 1: Recognition rates for HMMs

IV. PRACTICAL CAPABILITIES OF BIOMETRICS

The previous section offered a very concise review of our work in evaluating a range of biometric techniques for incorporation into a practical CE system to control and manage the recording and playback of video streams over a home network. However it does not address the practicalities of what is actually feasible for a real working system. We will summarize some of our current experiences in this section.

A. Practicalities of Fingerprint Acquisition and Analysis

The biggest issues with the use of fingerprints in CE systems continue to be the reliability of the acquisition process and the cost of the sensing technology. Capacitative sensors can provide adequate reliability but the user must pay attention to the acquisition process as too much or too little finger pressure will lead to poor image quality and authentication failure. We have also noted some reliability issues for these sensors in high usage scenarios and their costs remain too high for most CE applications.

An alternative is provided by *swipe sensors* where the finger is drawn across a row of scanning sensors. These 1D devices provide significant cost savings over 2D sensing technology but suffer from inaccuracies in measuring or determining the swipe speed during the acquisition process. The resulting raw fingerprint images can vary by as much as +/-20% along their length leading to inaccuracies in the subsequent pattern matching algorithms.

Once these acquisition issues are solved the actual enhancement of the acquired images and the reliable

extraction of *minutiae patterns* are tractable problems and can be readily implemented in standard embedded systems. Various pattern matching algorithms are available with good performance for the more mature techniques such as the ICP algorithm. It is worth commenting that most algorithms are designed for *forensic* matching of the extracted feature patterns and are not good at fast rejection of patterns. We believe that some improvements can be made in this context for CE applications and expect to present more concrete results in a future publication.

B. Practicalities of Facial Recognition

Facial recognition can provide good levels of authentication, particularly when applied to a continuous sequence of images. On single images it will be difficult to reliably achieve more than 85% accuracy in a consumer environment with uncontrolled pose and illumination [28]. However, if we acquire a continuous video sequence and accept that occasional positive recognition events provide enough confidence then face recognition becomes a very powerful tool for consumer applications.

One useful and non-intrusive approach is to mount a webcam on the TV and direct its field of view to cover the seating area for TV viewing. This *couch-spy* can be used to monitor the people viewing content and to take decisions based on continuous authentication. If the authentication criteria are chosen carefully this is a nice way to ensure that at least one genuine family member is watching protected content or recordings.

The other practical issue with the use of facial recognition techniques in consumer applications is that of enrollment. This is difficult if we try to use classical techniques such as PCA due to the complexity of the training process. It is also difficult for DCT techniques because we need to construct different, non-overlapping models to cover different pose and illumination conditions, although CLAHE normalization can help with the illumination issues. Fortunately we have had good results with HMM based techniques and it appears that capturing 2-3 separate face images at slightly different pose angles is sufficient to extrapolate reliable recognition models. These enrollment images can be easily captured from a single video stream allowing a single-step enrollment process. Details will be provided in a forthcoming publication.

C. Key Authentication Vs Key Generation

A core aspect of our system is the actual encryption of content streams. We employ standard hybrid encryption techniques which require the generation of local public/private key pairs [1]. Ideally we would like to be able to regenerate the private key solely from the biometric data as this would remove the need to permanently store the private key in digital form. However biometric data is notoriously fickle and known techniques to reliably generate keys directly from biometric data require rather complex implementations [33] which are unlikely to suit the computational power and resource availability of embedded CE appliances.

The alternative approach is to create keys using conventional random key-generation methods, associating each newly-created key with a biometric token. Thus, during the enrollment process for a new "identity" a corresponding public/private key pair is created and irrevocably associated with that identity. We refer to this approach as *key authentication* as opposed to the ideal case of *key generation* where the private key can be regenerated directly from some aspect of the biometric token. The latter enables content to be accessed on any system or device which supports the same *key generation* scheme as long as the appropriate biometric token is also provided.

Our initial tests have used the former approach, generating key pairs from a standard hashing algorithm. This does have the disadvantage that private keys need to be exchanged with remote systems in order to view recorded content on another playback device. A scheme to enable such key exchanges was reported and detailed in earlier work [1]. We can also report that some of our recent research, using local features from fingerprint patterns, appears to offer a promising approach to the *key generation* problem and we expect to report further on this in a forthcoming publication.

V. Stream Encryption and Playback

The problem of encrypting an MPEG video stream is, in itself, a somewhat non-trivial task. To understand this it is useful to consider the structure of an MPEG data stream – the case for MPEG-2 is illustrated in ***Fig 5*** below. The data is organized into a sequence of *frames* each of which consists of a set of horizontal *slices*. Each *slice* is further divided into square *macroblocks* comprising four *blocks* each of which is 8x8 pixels in size. This arrangement facilitates the encoding/decoding of video data using DCT based *codecs* techniques with the DCT methods being applied at the macroblock level.

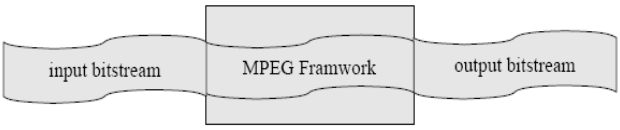

Fig 4: High level view of video stream encryption

However the situation is more complex because frames can be one of three types: *I-Frames* or key frames contain all of the original block level data for that frame; *P-Frames* contain DCT blocks and motion vectors which tell the decoder to deduce some of its block-level data from the previous frame in a video sequence; *B-Frames* contain motion vectors pointing to both the previous and next frames of the video sequence.

Conceptually we can view an MPEG stream as a set of raw DCT data sitting in a hierarchical data framework as shown in ***Fig 4***. Ideally we would like to preserve the data framework without altering it as this will leave the video structures intact and enable standard video decoders/players to manage and handle the video stream without generating undesirable errors. Thus our key goal must be to perform encryption/decryption

on the raw DCT data, rather than on the MPEG stream as a whole.

A. MPEG Stream Manipulation Framework

In order to allow new encryption techniques to be tested, a core element of this software is an MPEG1/2 stream manipulation framework. This framework, written in the Python programming language, provides classes for manipulating an MPEG elementary stream at a bit level and decoding of the MPEG bitstream syntax.

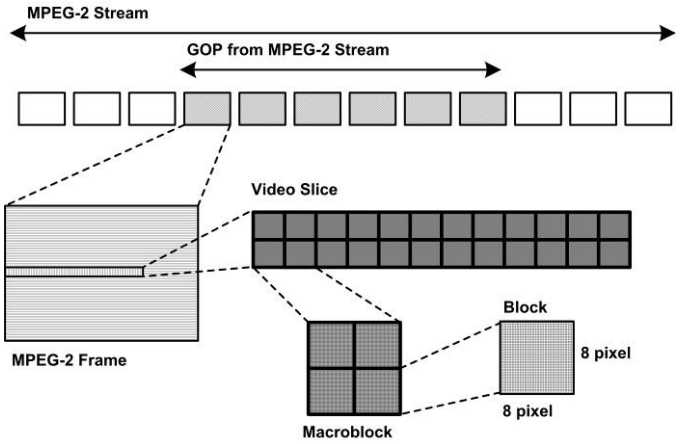

Fig 5: MPEG Elementary Stream Structure

The framework is not designed to be an MPEG decoder and does not provide inverse DCT operations. As the inverse DCT process is quite computationally intensive, it would not be practical to use a high level programming language such as Python to implement it. There are a number of freely available and very efficient MPEG1/2 codecs [18], which can be used if full stream decoding is required. However, for our test implementations we leave the full decoding of the video stream to a standard video player such as VLC [19].

This framework is suitable in situations where elements of an MPEG stream need to be changed or manipulated, without fully decoding the stream. The input is an MPEG-1/2 elementary stream and the output should be the same stream with some of the elements slightly changed. As these changes may result in an arbitrary change in the number of bits in the resultant stream, the framework must be able to add or remove any number of bits to the stream. The basic mode of operation of this framework is to read a stream, decoding each element of the MPEG bitstream (GOP, Frame, Slice, Macroblock, Block) and manipulating or encrypting some of these elements and subsequently outputting the modified stream.

The current version of our encryption algorithm works at the slice level of the MPEG stream. We change only the first four DCT data bytes in each video slice. When biometric key based encryption is used only these bytes are encrypted/decrypted – this reduces the chances of errors due to random data loss where the MPEG stream is transmitted over a network link. We also reduce this risk by starting encryption runs at the GOP level – thus each encryption (or decryption) run is restarted at the beginning of each GOP. If network errors occur and cause the loss of these encrypted bytes the start of the next GOP can still be readily identified and stream decryption can be restarted at the first data slice of that GOP. Where the bitstream is not encrypted we can simply swap the first two bytes in each video slice with the second pair of bytes to achieve a similar scrambling of the video stream. The latter technique is very useful for testing the main functionality of our bitstream manipulation framework.

An example of a scrambled video stream output is shown in ***Fig 6*** below. The original image was of our laboratory.

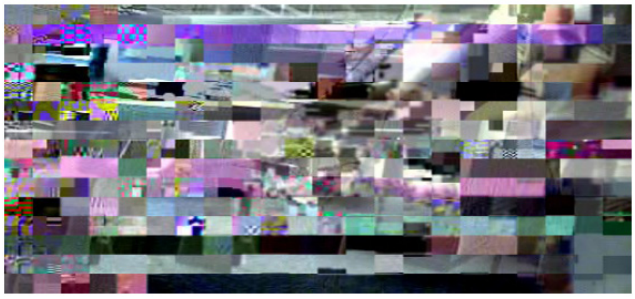

Fig 6: Example of Scrambled Video Stream Output

VI. TESTBENCH FOR BIOMETRIC IPTV

This is illustrated in ***Fig 7*** shown below. Device associations and control actions are managed by our universal middleware module [30] which is not shown here. Detailed discussions on building *metadevice representations* of networked home appliances are to be found in refs [29, 30]. Note that although we have found it convenient to integrate our testbed with our local middleware solution the techniques described in this paper are equally valid for any TCP/IP based content distribution system.

A. The Rebroadcast IPTV Server

This is a conventional desktop PC modified to act as a content acquisition and encoding platform. It is fitted with two TV-tuners equipped with MPEG-2 encoding hardware, which can be operated independently through various network services. Each TV rebroadcast uses about 4Mb of bandwidth to stream a DVD-quality MPEG-2 stream in real-time. Multiple users can access each TV rebroadcast but only the first user can change channels or activate other PVR services which are available on the server-side [31].

This server can also play and stream DVDs over the network and it offers two movie services which simulate *video-on-demand* (VOD) services which could be offered to home users. The first service streams movies via a *multicast* network address using UDP multicast with *forward error correction* (FEC) enhancements – users can only view the current movie joining it at its current playback point, corresponding to a multi-start VOD service. The second service is a TCP/IP based VOD service where users can select one of several movies which are stored on the server hard disk, but this requires a separate TCP/IP connection and video stream for each movie viewed.

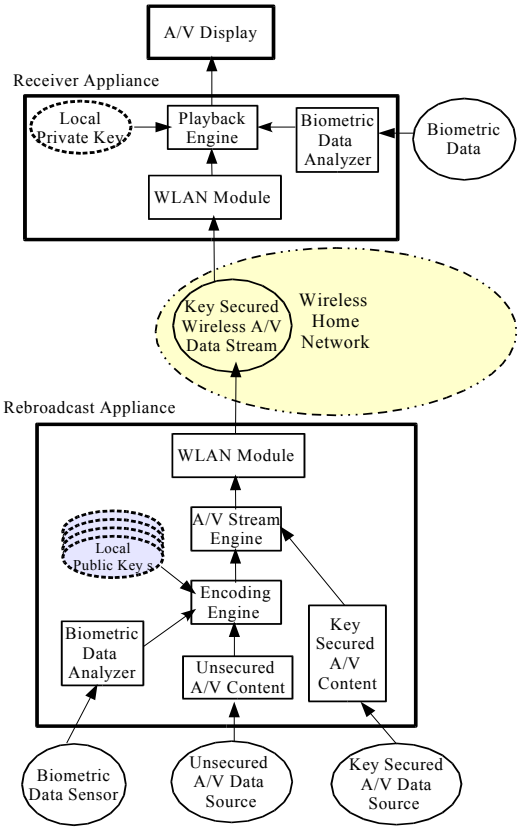

Fig 7: Biometric Content Encoding Subsystems

When operating at full capacity the rebroadcast server can reliably provide two TV stream rebroadcasts, one multi-user movie broadcast and one single user VOD rebroadcast on a conventional 802.11g wireless network. The availability of 4 distinct streaming channels should be adequate for most end-users, but additional channels and services could be added if an 802.11n or a combined 802.11a/g network was used. Employing MPEG-4 encoding for the movie services would further reduce the bandwidth requirements for the movie services.

B. The TV-Player Clients

Our clients are based on mini-ITX motherboards which provide PC-like compatibility in a small form factor suitable for use as set-top boxes. The system software is based on open-source operating system components and video players. An overlay is provided by a Python-based application which also manages interactions with the network middleware.

Users control the system using conventional A/V remote controls which operate an on-screen selection of the various networked services (TV1, TV2, multistart and VOD1). Where additional control functions are provided by a network service (e.g. for TV1 and TV2) the remote control switches to provide appropriate control functions such as switching channels on the TV tuner or activating a PVR recording function.

For encrypted services each player also features a USB fingerprint scanner and a video camera mounted on a local Plasma-TV screen. Thus, in addition to providing playback for streamed video content each TV-player also provides fingerprint authentication services and operates a continuous face-detection process which searches for the faces of people who are viewing the local TV screen. When faces are detected the TV-player captures and extracts the face regions and, depending on its operating mode it may request face recognition service from the IPTV rebroadcast server.

C. Networked PVR Appliance

As one of our key goals is to enable biometric encoding and decoding of content it made sense to integrate our PVR appliance [31] into the testbench. This appliance is network enabled and incorporates transcoding capabilities; it was originally designed for "harvesting" content from a broadband connection and redistributing this to other networked appliances over a home network. For this testbench it has been modified to act as a biometric encryption/decryption module and as a storage repository for secured content. Because our stream manipulation framework currently only supports MPEG-2 streams the PVR is also configured to only operate on MPEG-2 streams.

The main transcoding engine is implemented in Python which allows rapid prototyping of the transcoder workflow. Lower level components such as the A/V codecs and the A/V multiplexer and demultiplexer modules are built using open-source library components with Python or C/C++ wrappers as appropriate. The A/V buffer interface provides access to the RAM buffer extensions of the filesystem and is implemented independently from the transcoder module.

VII. Operational Aspects of Biometric Access Control

Our testbench is designed to explore various operational modes for biometric control of user access to digital content. If such modes of monitoring content are to offer workable alternatives to conventional DRM then they much be easy to use and should integrate easily with normal content viewing and recording activities. We can say that there are two principle operating modes for a user who is "sitting on the couch": either they will wish to view restricted content, or they may wish to record, or schedule for recording, some digital content. We outline below how our system can be configured to implement each of these modes.

A. Encrypted Broadcasting Mode

The rebroadcast IPTV server can be configured to operate in a continuous encryption mode. While operable in this mode it encrypts all content with the set of all public keys of authorized (enrolled) user from the local keystore. Client players require at least one of the corresponding private keys and the matching biometric token in order to be able to play back encrypted streams.

To access content an authorized user only needs to provide his fingerprint to the local scanner and, after a few seconds delay the local player should complete the authentication process and begin to decrypt the incoming content. To avoid loss of content we buffer the incoming stream until the authentication process is completed. In practice we enable continuous authentication from any enrolled user. Thus after initial fingerprint authentication is completed at a content

player it should not be necessary to re-authenticate as long as at least one enrolled user continues to view content provided by that player.

B. Encrypted Recording Mode

In a second operational mode the server can be configured to simply rebroadcast open content while content which a user wishes to record for future viewing is redirected to the PVR appliance. This redirection requires that a biometric authentication is provided before buffered content is saved to the PVR. Thus, the recording process is actuated by providing fingerprint data to the local client which transfers the authentication data to the main server. To simplify our implementation content is buffered and encrypted on the rebroadcast server and redirected to the PVR appliance where it is saved directly to the hard disk.

When playback is requested from the PVR the encrypted content is streamed directly to the playback client which requires a biometric token in order to begin decrypting the content.

C. Enrollment Toolset

Biometric data for enrollment can be input at the server-side, but it is more practical to enable enrollment from client-side appliances. There are slight risks that biometric data could be intercepted during the initial enrollment process, but in order to achieve widespread adoption of this type of content protection system we consider that ease-of-use should take priority. Thus we consider that local devices may be regarded as "trusted". Note that a secure mechanism for sharing private keys between appliances was described in [1] and requires that a user provide an identical biometric token at both appliances within a certain timeout period.

Two principle types of biometric data are used – active inputs and passive, or observational inputs. The active inputs, fingerprint data in our case, are entered from a laptop using a simple computer application and a fingerprint scanner. A user must provide 3-4 good quality fingerprints and the system will confirm enrollment. There is also a test function to verify that a user's fingerprint is correctly recognized by the system after enrollment. A similar application is used for face enrollment and requires that a standard USB webcam is connected to the laptop. We have found that HMM based face recognition is most reliable achieving reliable enrollment from a short video clip if the user simply nods their head slowly from left to right.

D. Observational Authentication

This mode was initially added to the system as an experiment. It provides an interesting alternative to single-event authentication and represents a new approach to controlling access to content. When observational authentication mode is activated the system scans a field of view in the vicinity of the living room viewing area and performs regular face detection on this viewing area. Detected face regions are compared against enrolled faces and as long as at least one of the people viewing the content is authorized the system continues to decode the media stream. If no authorized user is found the system enters a timeout mode and after a few minutes will flash a warning to the viewers. After a second timeout period it will flash a second warning and content decryption will cease about a minute later.

There are some interesting aspects to this approach. For example, when no faces are detected in the viewing area the system can automatically pause the video stream and, after a certain time interval it could optionally power down a display appliance. When faces are detected but none of these correspond to an authenticated user the system can either cease decoding the media stream, or alternatively can provide a reduced quality rendering of the content. When an authenticated user returns to the viewing area, or an authenticated fingerprint is input then normal viewing can be resumed.

VIII. SUMMARY & CONCLUSIONS

The goal of this research was to explore the use of biometric authentication as a means of controlling content access and management within a next-generation home networking environment where streamed content is the standard mode for providing A/V entertainment. We summarized several forms of biometric authentication, particularly fingerprint and facial recognition based systems. A summary of the practical capabilities and reliability of these technologies was also given.

We also described our networked testbed. Although this is based on our proprietary middleware infrastructure the techniques described in this paper are independent of the underlying implementation. We also introduced a distinction between *key authentication* and *key generation* systems and explained how the latter is still too unreliable and computation-intensive for CE applications.

Our overall conclusion is that biometric techniques are becoming sufficiently reliable and mature to warrant further exploration of their use in CE applications. There are still many problems to be solved but the potential to address fundamental issues such as access control to digital content is quite compelling.

REFERENCES

[1] P. Corcoran & A. Cucos, Techniques for securing multimedia content in consumer electronic appliances using biometric signatures, IEEE Transactions on Consumer Electronics, Vol. 51, No. 2, May 2005, pp. 545-551

[2] P. Corcoran, & G. Costache, Automated sorting of consumer image collections using face and peripheral region image classifiers, IEEE Transactions on Consumer Electronics, Vol. 51, No. 3, August 2005, pp. 747-754

[3] D. Maltoni, D. Maio, A.K. Jain, and S. Prabhakar, Handbook of Fingerprint Recognition; Springer, 2003

[4] S. Furui, An overview of speaker recognition technology in Proc. ESCA Workshop on Automatic Speaker Recognition Identification and Verification, 1994, pp. 1—9

[5] P. Corcoran, A. Cucos, & F. Callaly, IP-TV Stream Rebroadcasting Appliance for Home Use, submitted to IEEE Transactions on Consumer Electronics, July 2006.

[6] P. Corcoran, A. Cucos, and F. Callaly, A Universal Home MultiMedia Environment for Home Networking, at RoEduNet 2006, Sibiu, Romania; Volume 1, 1st-3rd June 2006 Page(s): 34-42

[7] P. Corcoran, C. Iancu and G. Costache, Improved HMM based Face Recognition System at OPTIM 2006, Brasov, Romania; Volume IV, 21-24 May 2006, Page(s): 143 - 146

[8] P. J. Besl and N. D. McKay. A method for registration of 3-d shapes. IEEE Trans. Pat. Anal. and Mach. Intel. 14(2), pp 239-256, Feb 1992.

[9] Y. Chen, G. G. Medioni. Object modelling by registration of multiple range images. Image and Vision Comp. 10(3), pp 145-155, 1992.

[10] C. Dorai, J. Weng, A. K. Jain. Optimal registration of object views using range data. IEEE Trans. Pat. Anal. and Mach. Intel. 19(10), pp 1131-1138, Oct 1997.

[11] T. Masuda, N. Yokoya. A robust method for registration and segmentation of multiple range images. Comp. Vision and Image Under. 61(3), pp 295-307, May 1995.

[12] G. C. Sharp, S. W. Lee, D. K. Wehe. Invariant features and the registration of rigid bodies. Proc. IEEE Int. Conf. on Robotics and Autom., pp 932-937, 1999.

[13] Z. Y. Zhang. Iterative point matching for registration of free-form curves and surfaces. Int. J. of Computer Vision, 13(2), pp 119-15, Oct. 1994.

[14] S. Rusinkiewicz and M. Levoy, Efficient Variants of the ICP Algorithm, Proceedings of the 3rd International Conference on 3D Digital Imaging and Modeling, jp145-152, Quebec, Canada, May 2001.

[15] H. Chui. Non-Rigid PointMatching: Algorithms, Extensions and Applications. PhD thesis, Yale University, 2001.

[16] H. Chui and A. Rangarajan. A new algorithm for non-rigid point matching. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2:44–51, 2000.

[17] Z. M. Kovacs-Vajna, A fingerprint verification system based on triangular matching and dynamic time warping, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol 22, No. 11, p. 1266-1276, Nov 2000.

[18] X. Jiang, and Y-W, Yau, Fingerprint minutae matching based on the local and global structures, Proceedings of 15th International Conference on Pattern Recognition, p 1038-1041, 2000.

[19] X. Chen, J. Tian, X. Yang, and Y. Zhang, An algorithm for distorted fingerprint matching based on local triangle feature set.

[20] X. Zhao, Y. Wang, J. Qi, and X. Zheng, Non-Alignment Fingerprint Matching Based on Local and Global Information, Proceedings of the First IEEE International Conference on Innovative Computing, Information and Control (ICICIC), 2006.

[21] J. Gu, J. Zhou, and C. Yang, Fingerprint Recognition by Combining Global Structure and Local Cues, IEEE Transactions on Image Processing, Vol 15, No 7, p. 1952-64, July 2006.

[22] A. N. Marana, and A. K. Jain, Ridge based fingerprint matching using Hough Transform, Proceedings of the XVIII Brazilian Symposium on Computer Graphics and Image Processing (SIBGRAPI '05), 2005.

[23] P. Corcoran, G. Costache, & R. Mulryan, Automatic indexing of consumer image collections using person recognition techniques, International Conference on Consumer Electronics, 2005 Digest of Technical Papers, p:127 - 128, Jan. 2005.

[24] P. Corcoran, M. Ionita, G. Costache, "Pose-invariant face recognition using AAMs", International Conference on Optimization of Electrical and Electronic Equipment, Brasov, Romania., May 2006

[25] S. Pizer, E. Amburn, J. Austin, R. Cromartie, A. Geselowitz, T. Greer, B. Romeny, J. Zimmerman, and K. Zuiderveld. Adaptive histogram equalization and its variations. Computer Vision, Graphics, and Image Processing, 39:355-368, 1987

[26] P. Corcoran, G. Costache , Automatic Person Retrieval from Image Collections, ESA-EUSC Conference, Frascati, Italy, Oct 2005

[27] P. Corcoran, G. Costache, R. Mulryan, and E. Steinberg, In-camera person-indexing of digital images, International Conference on Consumer Electronics, 2006. ICCE '06. 2006 Digest of Technical Papers, January 2006

[28] G. Costache, Advances in Automated Image Categorization: Sorting Images using Person Recognition Techniques PhD thesis, NUI, Galway, Jan 2006.

[29] P. Corcoran, F. Callaly, Rapid Prototyping of Networked A/V CE Appliances, The International Conference on Computer as a Tool, EUROCON 2005, Volume 2, Page(s):1312 - 1315, Belgrade, Serbia, Nov 2005..

[30] P. Corcoran, A. Cucos, and F. Callaly, Home Networking Middleware Infrastructure for Improved Audio/Video Appliance Functionality and Interoperability, The International Conference on Computer as a Tool, EUROCON 2005, Volume 2, Page(s):1316 - 1319, Belgrade, Serbia, Nov 2005..

[31] F. Callaly, P. Corcoran, Architecture of a PVR appliance with 'long-tail' Internet-TV capabilities, IEEE Transactions on Consumer Electronics, Volume 52, Issue 2, Page(s):454 – 459, May 2006

[32] P. Corcoran, C. Cucu, F. Callaly, and A. Cucos, Real-Time Fingerprint Analysis and Authentication for Embedded Appliances, IEEE International Conference on Consumer Electronics, Las Vegas, Jan 2007.

[33] W. Yong-Dong, Method of using biometric information for secret generation, published US Patent Application US 20040148509.

Peter Corcoran received the BAI (Electronic Engineering) and BA (Math's) degrees from Trinity College Dublin in 1984. He continued his studies at TCD and was awarded a Ph.D. for research work in the theory of Dielectric Liquids. In 1986 he was appointed to a lectureship in Electronic Engineering at NUI, Galway. His research interests include embedded systems, home networking, digital imaging and wireless networking technologies.

Alexandru Cucos received his B.S. degree in Electronic Engineering from "Transilvania" University from Brasov, Romania, in 1997. At the same university he received in 1998 M.S. degree in Electronic Design Automation. He received a M.Eng.Sc. degree in electronic engineering at National University of Ireland, Galway in 2001. Currently he is a senior research engineer working in the Consumer Electronics Research Group at National University of Ireland, Galway. His research interests include network streaming of multimedia content, embedded systems design, communication network protocols, and biometric sensing techniques.

Frank Callaly received the B.Eng. (Electronic & Computer Engineering) degree from the National University of Ireland, Galway in 2003. He continued his studies with the Consumer Electronics Research Group at the National University of Ireland, Galway and is currently pursuing a Ph.D. degree in Electronic Engineering. His current research interests include digital video encoding, multimedia delivery techniques and home networking.

Claudia Iancu received the B.S degree in Communications and M.S. degree in Signal Processing from 'Politehnica' University of Bucharest, Romania in 2003 and 2004 respectively. She is currently pursuing a PhD. degree in Image processing. Her current research interests include signal processing and pattern recognition with applications in face recognition.