

채팅, 화상채팅의 실태 분석 및 기술적 해결 방안의 제안(II)

- 기술적 해결 방안을 중심으로 -

Proposal of Technical Method for Solving Internet Harmful Contents

조동욱, 정진우, 한선아, 전환규, 정인호*
충북과학대학 정보통신학과*

Cho Dong-Uk, Chung Jin-Woo, Han Ah-Sun,
Chun Gyu-Hwan, Chung In-Ho*
Chungbuk Provincial University of
Science & Technology*

요약

본 논문에서는 채팅, 화상 채팅 사이트가 가지고 있는 문제를 기술적으로 해결하기 위한 방법을 제안하고자 한다. 이를 위해 성인 인증 기능의 강화를 위한 얼굴 생체 인증, 그리고 음란 행위를 검출하여 차단하기 위한 영상 신호 처리와 음향 신호 처리 등에 대해 제안하고자 한다. 끝으로 실험에 의해 제안한 방법의 유용성을 입증하고자 한다.

Abstract

This research proposes the technical solving method for blocking Internet harmful sites such as chatting and image chatting. For this, biometric verification using human front-faces is proposed. Also, signal processing method is proposed for blocking pornographic acts. Finally, the effectiveness of this paper is demonstrated by experiments.

I. 서론

인터넷 유해콘텐츠란 통상 음란 동영상, 자살 사이트, 병역기피사이트, 자퇴사이트, 채팅과 화상 채팅 사이트 등을 일컬으며 이를 인터넷 역기능의 대표적인 형태로 분류하고 있다. 그러나 현재 인터넷 역기능중 가장 문제가 심각한 것은 채팅과 화상 채팅으로 여겨지고 있다. 원래 채팅이란 자판, 음성, 영상 등을 통해 실시간으로 대화를 나누어 자료 공유와 의견 교환을 행하는 보편적 서비스를 창하였다. 천리안 등과 같은 PC통신에서 처음 서비스를 제공했으며 인터넷에서 세이클럽, 하늘사랑, 지오피아, 클럽 5678 등이 다양한 채팅서비스를 제공하고 있다 [1]~[9]. 채팅의 형태는 여러 사람이 함께 이야기하는 형태, 일대일로 대화하는 형태, 쪽지를 주고받는 형태, 상대방의 얼굴

을 보며 이야기하는 형태 등 여러 형태가 있다. 그러나 현재 인터넷상에서 채팅은 건전한 만남이나 의견 교환의 장이 아니라 불륜이나 불건전한 만남의 도구로 쓰이는 즉, 순기능보다는 역기능이 더욱 문제가 되는 상황이 되었다. 그전에는 채팅을 통해 미·기혼자나 배우자가 없는 사람들이 사이버 성희롱을 주로 하여 그 피해가 사이버 상에 머물렀지만 이제는 사이버 성희롱이라는 용어가 없어질 정도이며 채팅을 통해 실제 오프라인에서 만나 성 범죄와 성타락이 행해지고 있는 실정이다. 이제는 채팅을 통해 번개로 만나 불륜을 저지르고 어린 학생들이 원조 교제를 행하는 그 심각성이 너무 크고 방대한 실정이 되어 버렸다 [10], [11]. 심지어 채팅사이트를 운영하는 업자 입장에서 이의 문제점을 이미 충분히 인지하고 있는

상황에서도 접속자가 많아 이윤 창출이 많이 되는 관계로 이를 조장하거나 방치하는 경우가 대부분이기 때문에 이의 심각성 및 해악성은 반드시 거론하고 가야 하는 실정이었다. 아니 어찌 보면 이를 조장할 수 있는 메뉴를 많이 만들어 접속자수를 늘리는데 더 많은 주안점을 두고 있는 것이 아닌가하는 의구심까지 들게 하는 것이 현 채팅사이트들이 가지고 있는 실정이다.

이를 방지할 경우 가장 큰 문제점은 이미 언급한 바와 같이 원조교제와 주부 탈선 문제 등이며 근래는 화상 채팅을 통해 남녀간의 성기 노출과 자위 행위 등을 보여주는 등 그 폐해가 더 이상 그대로 방치할 수 없는 실정에 도달한 것이 현재 상황이다. 특히 화상 채팅으로 인해 청소년들이 채팅방에서 남녀의 성기를 드러낸다는 것은 상당한 사회적 문제가 되고 있으며 그 가입자 수가 몇 백만 명에 이르는 등 실로 심각한 상황이 아닐 수 없다. 따라서 본 논문에서는 이와 같은 채팅, 화상 채팅의 문제점을 기술적으로 해결하기 위한 방법을 제안하고자 한다. 이를 위해 성인 인증 기능의 강화 즉, 아이디와 패스워드뿐 아니라 성인 인증에 필요한 얼굴 인증까지 행하여 미성년자들이 아는 사람의 주민 번호 도용이나 주민 번호 생성에 의해 채팅, 화상 채팅 사이트에 가입하는 것을 원천적으로 차단하고자 한다. 아울러 실제의 채팅 시 화상 캠을 통해 성기를 노출시키는 행위나 음란성 소리 등을 내는 것을 기술적으로 차단하고자 한다. 끝으로 실험에 의해 제안방법의 유용성을 입증하고자 한다.

2. 채팅, 화상 채팅에 대한 문제점 해결 방안

채팅, 화상채팅, 음란 콘텐츠에 대한 문제점을 해결하기 위해 다음과 같은 방안이 강구되어야 한다.

첫째, 채팅, 화상 채팅 사이트 운영자가 어느 정도는 공익을 고려하여 운영을 해야 한다. 가입자 수만

많으면 수입이 많다는 식으로 무방위 상태로 채팅의 문제점을 방조하는 것은 사회에 크나 큰 죄악을 짓는다는 생각을 가져야 할 것이다. 심지어 앞장서서 불건전한 만남을 유도할 수 있는 메뉴를 개설해 두고 있는데 이를 금지시켜야 한다. 또한 화상 채팅방에서 투명인간 기능을 약 1,500원 정도에 구매할 수 있게 하여 수많은 가입자들로부터 엄청난 이득을 챙긴 채팅방 운영자도 있는 실정이므로 운영자들에 대한 도덕성 회복도 시급한 과제이다.

둘째, 특정 단어 입력을 못하도록 해야 한다. 예를 들어 채팅상에서 불건전한 만남을 유도할 수 있는 단어 등에 대해 기술적으로 필터링을 할 수 있는 기능 등을 첨부해야 한다.

셋째, 네티즌들이 몰리는 새벽 시간대에 비공개방을 아예 폐쇄해야 한다.

넷째, 성인 인증 기능의 강화이다. 현재는 주민번호만 맞으면 바로 아이디와 패스워드가 원하는 대로 주어지기 때문에 이를 얼굴 인증 등을 첨부하여 청소년들이 성인 채팅 사이트, 성인 화상 채팅 사이트에 접속하는 것을 원천적으로 차단해야 한다.

다섯째, 성인의 경우라도 화상 채팅시 성기를 그대로 노출하거나 자위행위를 하는 행위 그리고 교성에 가까운 소리 등을 내는 것은 원하는 사람의 경우 차단할 수 있는 기술적 차단 방법이 강구되어야 한다.

여섯째, 현재 채팅사이트, 화상 채팅사이트, 음란 콘텐츠 안내, 폰팅 등이 대부분 스팸 메일을 통해 알게 되는 경우가 많으므로 스팸의 원인이 되는 이메일 추출기의 기능을 정지시킬 수 있는 방법이 강구되어야 하리라 여겨진다.

일곱째, 대학에서도 IT학과의 경우 기능 교육과 이론 교육만이 이루어지고 있는데 이와 더불어 정보통신윤리교육도 행해져야 하리라 여겨진다. 각종 바이러스 퇴치 알고리즘 개발에서부터 음란 유해 콘텐츠 차단법등 다양한 형태의 정보 통신 윤리 교육이 이루어져야 하리라 여겨진다.

3. 기술적 해결 방안의 제시

3.1 성인 인증 기능의 강화

현재 PC에서는 화상 채팅을 위한 저가의 웹캠이 설치되어 있다. 19세 이상이 가입하는 채팅사이트에 아는 사람의 성인 주민번호로 가입하는 경우가 많은데 이를 위해 아이디, 패스워드 외에 성인인증을 위한 얼굴 생체 인증을 행해야 하리라 여겨진다. 이를 위해 얼굴형과 얼굴 특징벡터를 추출하여 이를 행하고자 하며 아래 표 1에 얼굴형의 예를 그리고 그림 1에 얼굴내 특징 추출 벡터를 나타내었다. 또한 표1에서 구한 특징값을 최대값과 최소값의 차를 20등급으로 나누어 정규화를 행해 최종적으로 절대치의 합을 통해 본인인지의 인증을 행하고자 한다.

[표 1] 얼굴 생체 인증을 위한 얼굴형의 예

대표적 얼굴형
역삼각형
동근형
네모형
계란형
삼각형
마름모형



특징 요소	내용(거리)
f1	머리, 이마
f2	이마, 눈썹
f3	눈썹, 눈
f4	눈, 코
f5	코, 입
f6	입, 턱
f7	코, 코
f8	입, 입
f9	좌입, 좌측
f10	우입, 우측
f11	좌측, 우측
f12	코중앙, 좌측
f13	코중앙, 우측

▶▶ 그림 1. 얼굴내 생체 인증 특징 벡터의 선정

3.2 음란 영상 차단

3.2.1 영상분할

영상분할을 위해 텍스처 분석을 행한다. 본 논문에서는 텍스처 분석을 통계적 방법인 GLCM 방법중 변화에 민감하게 반응하는 모멘트와 동일 성질에 반응하는 동차성을 조합하는 방법을 참조하여 빈도수를 통한 GLCM내 의 값들의 정규화와 동차성의 정규화를 행하는 방법을 제안하여 신체의 주요 성적 부분의 영역을 추출하고자 한다. 우선 GLCM(Gray Level Co-occurrence Matrix) $G[i, j]$ 는 변위벡터 $d = (dx, dy)$ 를 결정하고 명도 i 와 j 를 갖고 d 만큼 떨어져 있는 화소들의 쌍을 계산하여 생성한다. 이때 전체 빈도수를 GLCM의 각 원소값들에 나누어 줌으로써 $\sum G[i, j] = 1$ 이 되도록 조정한다. 따라서 이를 이용하여 모멘트 M 은 하식 (1)과 같이 구한다.

$$M = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} (i-j)^2 G[i, j] \quad (1)$$

또한 정규화된 동차성 H 는 하식 (2)와 같이 새로 이 정의하여 그 값을 구한다.

$$H = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \frac{255 - |i-j|}{255} G[i, j] \quad (2)$$

이를 4×4 윈도우 크기로 연속 수행하여 두 개의 배열 $[A]$ 와 $[B]$ 에 각각 저장한다. 이때 배열 $[A]$ 와 배열 $[B]$ 에 대해 특정한 임계치의 값으로 임계치 작업을 행하며 이를 통해 배열 $[A]$ 와 배열 $[B]$ 에 대해 OR 작업을 수행한다.

이의 결과는 신체의 주요 성적 부위에 해당하는 영역이 되며 이를 Y 축으로 프로젝션을 행하여 차후 형성된 히스토그램에 대한 분석을 행한다.

3.2.2 히스토그램 분석과 곡선 적합

이제 앞절에서 구한 히스토그램을 분석하여 곡선 적합을 수행하고자 한다. 곡선 적합시 여성과 남성의 신체의 주요 성적 부위는 2차곡선이나 일량분포로 나타낼 수 있으므로 이에 대한 곡선적합을 수행한다. 우선 $(M+1)$ 개의자료

$\{(x_i, y_i) \mid i = 0, 1, \dots, M\}$ 를 $n (< M)$ 차의 다항식 $P_n(x) = \sum_{k=0}^n a_k x^k$ 에 근사시키는 곡선적합은 아래와 같다.

$$E = \sum_{i=0}^M (y_i - P(x_i))^2 \quad (3)$$

위 식을 최소화하는 a_0, a_1, \dots, a_n 을 구하기 위해서는 (3)식에서

$$\begin{aligned} E &= \sum_{i=0}^M y_i^2 - 2 \sum_{i=0}^M P(x_i) y_i + \sum_{i=0}^M (P(x_i))^2 \\ &= \sum_{i=0}^M y_i^2 - 2 \sum_{i=0}^M \left(\sum_{j=0}^n a_j x_i^j \right) y_i + \sum_{i=0}^M \left(\sum_{j=0}^n a_j x_i^j \right)^2 \\ &= \sum_{i=0}^M y_i^2 - 2 \sum_{j=0}^n a_j \left(\sum_{i=0}^M y_i x_i^j \right) + \sum_{j=0}^n \sum_{k=0}^n a_j a_k \left(\sum_{i=0}^M x_i^{j+k} \right) \end{aligned} \quad (4)$$

(4)식을 최소화하려면 $j = 0, 1, \dots, n$ 에 대하여

$$\frac{\partial E}{\partial a_j} = (-2 \sum_{i=0}^M y_i x_i^j) + 2 \sum_{k=0}^n a_k \sum_{i=0}^M x_i^{j+k} = 0 \quad (5)$$

따라서 $(n+1)$ 개의 미지수 a_j 에 대하여

$$\sum_{k=0}^n a_k \sum_{i=0}^M x_i^{j+k} = \sum_{i=0}^M y_i x_i^j \quad (6)$$

이차곡선에 대해서는

$$a_0 \sum_{i=0}^M x_i^0 + a_1 \sum_{i=0}^M x_i^1 + a_2 \sum_{i=0}^M x_i^2 = \sum_{i=0}^M y_i x_i^0 \quad (7)$$

$$a_0 \sum_{i=0}^M x_i^1 + a_1 \sum_{i=0}^M x_i^2 + a_2 \sum_{i=0}^M x_i^3 = \sum_{i=0}^M y_i x_i^1 \quad (8)$$

$$a_0 \sum_{i=0}^M x_i^2 + a_1 \sum_{i=0}^M x_i^3 + a_2 \sum_{i=0}^M x_i^4 = \sum_{i=0}^M y_i x_i^2 \quad (9)$$

따라서 a_0 는 아래와 같이 구할 수 있다.

$$a_0 = \frac{\begin{vmatrix} \sum y_i x_i^0 & \sum x_i^1 & \sum x_i^2 \\ \sum y_i x_i^1 & \sum x_i^2 & \sum x_i^3 \\ \sum y_i x_i^2 & \sum x_i^3 & \sum x_i^4 \end{vmatrix}}{\begin{vmatrix} \sum x_i^0 & \sum x_i^1 & \sum x_i^2 \\ \sum x_i^1 & \sum x_i^2 & \sum x_i^3 \\ \sum x_i^2 & \sum x_i^3 & \sum x_i^4 \end{vmatrix}} = \frac{|D_0|}{|D|} \quad (10)$$

마찬가지로 a_1, a_2 로 (11), (12)것과 같이 하여 구할 수 있게 된다.

$$a_1 = \frac{|D_1|}{|D|} \quad (11)$$

$$a_2 = \frac{|D_2|}{|D|} \quad (12)$$

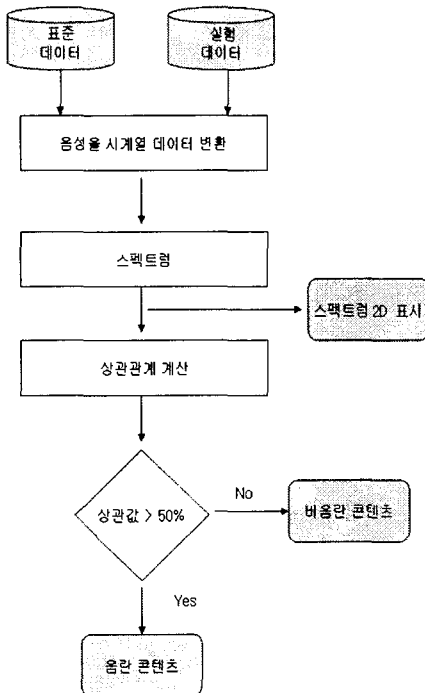
또한 일량분포에 대한 곡선적합은 $y = a + bx$

에 대해 아래식과 같이 정리되어 일량분포에 대한 곡선적합을 수행할 수 있게 된다.

$$\begin{bmatrix} a \\ b \end{bmatrix} = \frac{1}{M \sum x_k^2 - (\sum x_k)^2} \begin{bmatrix} \sum x_k^2 & -\sum x_k \\ -\sum x_k & M \end{bmatrix} \begin{bmatrix} \sum y_k \\ \sum x_k y_k \end{bmatrix} \quad (13)$$

3.3 음향 콘텐츠의 처리

이제 음란 콘텐츠에서 성행위시 발생하는 음향 콘텐츠를 처리하는 방법에 대해 제안하고자 한다. 이는 우선 음란사이트에서 나오는 성교시 신음 음향을 시계열 데이터로 변환하고 이를 2차원 스펙트럼으로 나타낸다. 차후 2차원 스펙트럼으로부터 특징을 추출하여 이를 표준 음향 신호와 상관 계수를 계산함으로써 음란 콘텐츠에서 음향 신호에 의한 음란사이트 차단을 행하고자 한다. 아래 그림 2.에 음향신호처리에 대한 전체 흐름도를 나타내었다.



▶▶ 그림 2. 음향신호처리에 의한 음란 사이트 차단 의 전체 흐름도

주어진 음란 콘텐츠에서 음향신호를 처리하기 위해서는 표준이나 실험 음향 신호로부터 이의 신호를 시계열 데이터로 변환하여 음향신호의 지문이랄 수 있는 스펙트럼을 생성해야 한다. 이를 위해 프라트 [12] 라 불리는 음향 신호처리 툴을 적용하고자 한다. 프라트는 암스테르담대학의 Paul Boersma와 David Weenink가 만든 음성 분석 및 변형 프로그램 패키지인데 이것은 인터넷 홈페이지에서 무료로 다운받을 수 있다. 작업 환경은 프라트 객체창과 프라트 그림창으로 나뉘어 지는데 기본적으로 프라트 객체창은 모든 분석 작업을 통제하는 통제실과 같은 역할을 하고, 프라트 그림창은 분석 결과를 그림으로 그리고 동시에 필요한 문자와 화살표, 선등으로 그려 넣은 뒤 그림 파일을 실을수 있는 기능을 가지고 있다. 표준 스펙트럼 신호의 특징값과 실험 대상 스펙트럼 신호의 특징값의 상관 계수를 계산해야 해당 사이트가 음란 콘텐츠인지 아닌지를 계산할 수 있다. 이를 위해 상관 값을 계산해 주어야 하는데 예를 들어 A신호와 B신호사이의 상관계수 값은 하식과 같이하여 구한다.

$$r = \frac{\sum_m \sum_n (A_{mn} - \bar{A})(B_{mn} - \bar{B})}{\sqrt{(\sum_m \sum_n (A_{mn} - \bar{A})^2)(\sum_m \sum_n (B_{mn} - \bar{B})^2)}}$$

4. 실험 및 고찰

4.1 영상 콘텐츠에 대한 실험 결과 및 고찰

영상 콘텐츠에 대한 실험은 IBM-PC 상에서 C언어를 이용하여 행하였다. 우선 그림 3이 여성이 화상 채팅하면서 성적 부위를 노출 시킨 장면이고 그림 4가 이에 대한 영역 분할 결과, 그림 5가 히스토그램 분석 결과이다. 마찬가지로 그림 6이 여성이 화상 채팅하면서 자위를 하는 장면을 나타낸 영상, 그리고 이에 대한 영역 분할 결과가 그림 7 이며 히스토그램 분석 결과가 각각 그림 8이다. 실험 결과에서 알 수

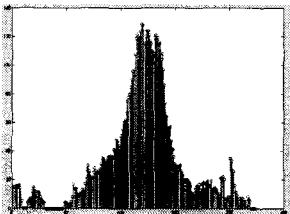
있듯이 본 논문에서 제안한 방법으로 화상 채팅시 보이는 음란 행위에 대한 영상을 정확히 처리할 수 있음을 확인할 수 있다.



▶▶ 그림 3. 화상 채팅 문제 장면



▶▶ 그림 4. 영역 분할 결과



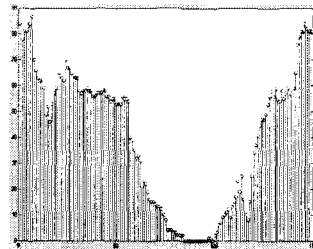
▶▶ 그림 5. 히스토그램(투영)



▶▶ 그림 6. 화상 채팅 문제 장면



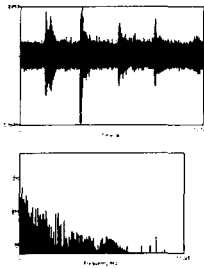
▶▶ 그림 7. 영역 분할 결과



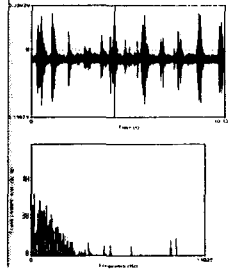
▶▶ 그림 8. 히스토그램(투영)

4.2 음향 콘텐츠에 대한 실험 및 고찰

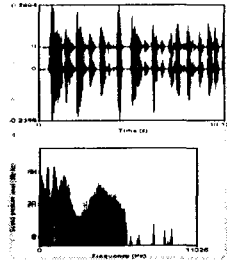
음향 콘텐츠에 대한 실험은 IBM-PC 상에서 C++을 사용하여 행하였다. 우선 인터넷 사이트에서 주로 많이 나오는 10개의 표준이 되는 실험샘플들을 wav 파일 포맷으로 변환하였다. 아래 그림 9~그림 18에 10개의 샘플에 대한 음향신호와 이에 대한 스펙트럼을 나타내었으며 스펙트럼 생성은 앞서 언급한바와 같이 프라트(Parrat)라고 하는 소프트웨어 툴을 이용하였다. 실험 결과는 표 2, 표 3, 표 4에서 알 수 있듯이 표준 샘플과 유사한 실험 대상(샘플 6과 샘플 7, 그리고 샘플 3) 음란콘텐츠의 음향신호를 분석하여 상관계수를 계산한 결과는 상관 계수값이 높게 나오는 음란 이라는 것을 정확히 지적해 주는 실험 결과임을 알 수 있었다. 향후 보다 다양한 표준 샘플을 선정하여 각종 음란 음향신호에 대한 효율적으로 차단이 이루어지도록 하기 위한 연구가 지속되어야 하리라 여겨진다.



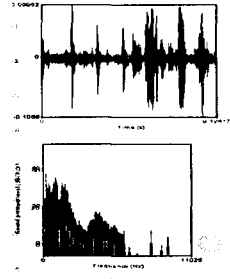
▶▶ 그림 9. 샘플1



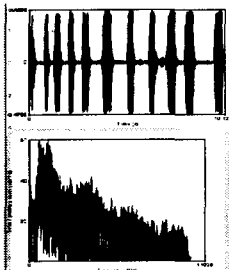
▶▶ 그림 10. 샘플2



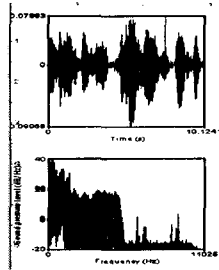
▶▶ 그림 17. 샘플9



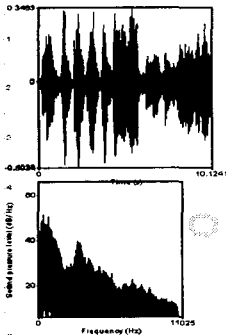
▶▶ 그림 18. 샘플10



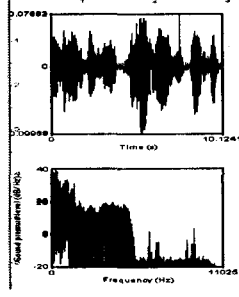
▶▶ 그림 11. 샘플3



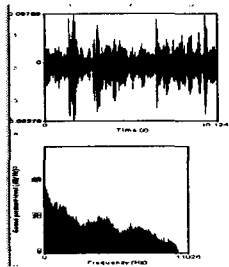
▶▶ 그림 12. 샘플4



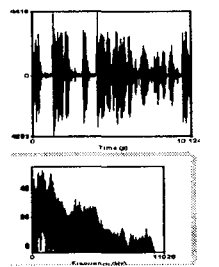
▶▶ 그림 13. 샘플5



▶▶ 그림 14. 샘플6



▶▶ 그림 15. 샘플7



▶▶ 그림 16. 샘플8

[표 2] 실험결과 (샘플 6과 실험대상 A)

	샘플6	실험대상A
샘플6	1.0000	0.5236
실험대상A	0.5236	1.000

[표 3] 실험결과 (샘플 7과 실험대상 B)

	샘플7	실험대상B
샘플7	1.0000	0.5660
실험대상B	0.5660	1.000

[표 4] 실험결과 (샘플 3과 실험대상 C)

	샘플3	실험대상C
샘플3	1.0000	0.9285
실험대상C	0.9285	1.0000

5. 결론

본 논문에서는 인터넷역기능의 대표적인 채팅, 화상 채팅사이트들이 가지고 있는 역기능에 대한 문제를 해결하기 위한 기술적 해결 방안을 제안하였다. 이를 위해 우선적으로 청소년 보호를 위해 성인 인증을 강화하였으며 얼굴 생체 인증을 행하였다. 얼굴 생체 인증을 위해 얼굴형에 대한 정보 그리고 얼굴내의 각 얼굴 요소들간의 특징 벡터를 추출하였다. 또한 텍스춰 분석, 곡선적합 등을 통해 음란 영상에 대

한 차단을 행하고, 음향 신호를 처리하여 음란 음향 신호를 차단하였다. 향후 보다 다양한 환경 하에서의 실험 수행 그리고 이를 상용화하여 인터넷의 사회적 역기능이 최소화되기 위한 지속적 노력의 경주 특히 청소년들을 인터넷 역기능으로부터 보호하기 위한 노력이 더욱 더 행해져야 하리라 여겨진다.

■ 참고문헌 ■

- [1] 세이클럽(<http://www.sayclub.com>)
- [2] 하늘사랑(<http://www.skylove.com>)
- [3] 지오피아(<http://www.geopia.com>)
- [4] 클럽 5678(<http://www.club5678.com>)
- [5] 아이미팅(<http://www.eyemeeting.co.kr>)
- [6] 오마이러브(<http://www.ohmylove.co.kr>)
- [7] 조이천사(<http://www.joy1004.co.kr>)
- [8] 컬러팅(<http://www.colorting.co.kr>)
- [9] 세이러브(<http://www.saylove.co.kr>)
- [10] 조선일보(<http://www.chosun.com>)
- [11] 조동욱, “채팅하면 생각나는 그 사람”, 충청일보 특별기고, 2003년 11월 17일
- [12] http://www.fon.hum.uva.nl/pratt/download_win.html