JDCS
디지털콘텐츠학회논문지
Journal of Digital Contents Society
Vol. 20, No. 12, pp. 2515-2524, Dec. 2019

음성과 영상 데이터 Key 를 사용한 이중 암호화 및 실시간 지능형 복호화

이 덕 규[1] · 엄 진 섭[2*]
[1]강원대학교 전자공학과 석박사 통합 과정
[2]강원대학교 전자공학과 교수

Dual Encryption and real-time intelligent Decryption using Speech and Image Data Keys

Duk-Kyu Lee[1] · Jinseob Eom[2*]
[1,2]Department of Electronics Engineering, Kangwon National University, Kangwon, Kerea

[요 약]

본 논문에서는 생체 데이터인 안면 데이터와 음성 데이터를 Key로 사용하여 이중으로 암호화 및 복호화를 수행하는 새로운 알고리즘을 제안하였고 LabVIEW 프로그램을 사용하여 이를 실시간으로 구현하였다. 암호화는 입력 데이터와 2개의 Key 데이터들을 각각 푸리에 변환한 후에 차례로 곱 연산을 수행하여 진행된다. 복호화는 두 단계로 진행된다. 우선 동일인 확인을 위하여 새로 수집된 2개의 Key 데이터들과 암호화에서 사용되었던 Key 데이터들 간의 유사성을 판별한다. 특히 음성 Key 데이터에 대해서는 웨이블릿 변환을 사용하여 특징 값들을 추출하였고 이들에 기계학습 ANN(Artificial Nerual Network)을 적용하여 판별의 정확성을 높였다. 다음으로 유사성이 모두 인정되는 경우에만 암호화 과정의 역순으로 연산하여 입력 데이터를 복원한다. 실험을 통하여 5초 이내에 복호화 과정이 완료될 수 있음을 확인하였다. 제안된 방식은 실시간으로 수집한 안면 및 음성 데이터 즉 2개의 동적인 생체 데이터를 Key로 사용하였다는 점에서 정적인 1개의 Key 만을 사용하는 기존의 방식보다 더욱 높은 보안성을 제공할 수 있다.

[Abstract]

In this paper, double encryption and decryption algorithm which uses facial data and voice data as keys was proposed and implemented in real time through LabVIEW programming. Encryption is a sequential multiplication process for each Fourier transformed data on an input data and two key data. Decryption is performed in two steps. First, the similarity between two newly captured key data and two key data already used in encryption is checked for an identical examination. Specially in case of voice key data, Wavelet Transform was used to extract features of data, then machine learning ANN was applied to those features to improve the accuracy of discrimination. Next, only when all similarities are recognized, the input data is restored through the reverse process of encryption. The experiment showed that the run-time for decryption is under 5 second. Compared to the existing scheme only with one static key, the proposed can provide higher security in that facial and voice data captured in real time, that is, two dynamic biometric data are used as keys.

색인어 : 암호화, 복호화, 푸리에 변환, 웨이블릿 변환, 기계학습, 랩뷰
Key word : Encryption, Decryption, Fourier Transform, Wavelet Transform, Machine Learning, LabVIEW

http://dx.doi.org/10.9728/dcs.2019.20.12.2515

Received 20 October 2019;**Revised** 30 November 2019
Accepted 15 December 2019

***Corresponding Author; Jinseob Eom**
Tel: ※ 개인정보 표시제한
E-mail: jeom@kangwon.ac.kr

Ⅰ. 서 론

21세기 정보화시대에는 인터넷이라고 하는 정보처리체계가 우리 생활의 중심에 놓여 있다. 인터넷이 우리 생활의 한 부분이 되어 많은 영향을 끼치게 된 것은 PC(Personal Computer)의 보급률이 높아지고 어느 곳에서나 쉽게 인터넷에 접속 할 수 있게 된 결과라 할 수 있다. [1]

과거부터 현재까지 PC 의 기하급수적인 보급과 발달로 인하여 다른 사람으로부터 자신의 정보 및 더 나아가 국가적인 기밀 등을 지킬 수 있는 보안의 필요성이 대두 되었고 그 결과로서 암호화 알고리즘이 발달하게 되었다. 이와 더불어 암호화 알고리즘을 해제할 수 있는 복호화 알고리즘 또한 급속도로 발달하게 되었다. 자신의 정보를 보호하려는 노력은 암호 작성법의 연구 동기가 된 반면에, 상대의 정보를 가로채려는 노력은 암호 해독의 연구 동기가 되었다. 암호학은 암호 작성과 암호 해독 양쪽을 포함하는 학문으로서 오늘날의 정보화시대에 매우 중요한 연구 및 응용분야이다[2].

암호화 기술 만큼이나 빠르게 진화하는 복호화 기술에 대비하기 위하여 이전보다 더욱 보안성이 높은 새로운 암호화 알고리즘이 요청되고 있는 실정이다. 본 논문에서는 푸리에 변환을 활용한 데이터들의 합성으로 암호화를 수행하고 또 이의 역 과정을 통하여 복호화를 수행하는 알고리즘을 제안하였으며 LabVIEW 프로그램을 사용하여 이를 구현하였다. 특히 암호화 및 복호화에 사용될 Key(비밀번호)로서 두개의 동적 Key(음성과 안면 실시간 생체 데이터)를 사용하였다. 이를 통하여 일반적으로 한 개의 정적인 Key 를 사용하는 기존의 암호화 알고리즘에 비하여 더 높은 보안성을 기대할 수 있다. 복호화는 두 단계로 진행되는데 처음에는 동일인임을 확인하기 위하여 실시간으로 새로 수집된 생체 Key 데이터들과 암호화에서 사용된 Key 데이터들간의 유사성을 판별한다. 이를 위하여 웨이블릿 변환 및 기계학습 ANN을 사용하여 판별의 정확성을 높였다. 또한 본 논문에서는 LabVIEW 프로그램을 사용하여 암호화 및 복호화 과정을 실시간으로 모니터링할 수 있도록 하였다.

본 논문은 다음과 같이 구성된다. 2장에서는 제안된 암호화 및 복호화 알고리즘에서 사용되는 기술들의 이론을 살펴보았다. 3장에서는 제안된 알고리즘에 대한 설명 및 이의 구현 과정을 상세히 기술하였다. 4장에서는 LabVIEW로 구현된 프로그램으로 암호화 및 복호화 실험을 수행하였고 사용된 데이터들과 실험 결과를 실시간 모니터링 창을 통하여 나타내었다. 5장에서는 전체적인 결론을 맺었다.

Ⅱ. 사용 프로그램 및 관련 이론

2-1 LabVIEW

LabVIEW는 Laboratory Virtual Instrument Engineering Workbench 를 줄인 말이다. 이것은 마이크로소프트 윈도우즈, 애플 매킨토시, 선 스파이크스테이션을 이용하는 PC 에서 운영될 수 있으며 강력하고 유연성 있는 해석 소프트웨어 시스템의 일종이다. LabVIEW 는 여러 범용의 C 또는 BASIC 개발 시스템처럼 프로그램 개발 도구이다. 다른 프로그래밍 시스템들이 코드를 생성하기 위하여 텍스트 기반의 언어를 사용하는 데 반하여 LabVIEW는 여러 가지 문법적인 사항들을 없애고 블록 다이어그램이라 불리는 흐름도를 사용하여 프로그램을 생성하는데, 우리는 이 언어를 그래피컬 프로그래밍이라고 부르며 간단히 G 언어라고 한다[3].

2-2 푸리에 변환(Fourier Transform)

본 논문에서는 푸리에 변환을 이용하는 새로운 암호화 알고리즘을 제안하였다. 시간 t와 더불어 변동하는 양 f(t)가 있다고 하면 이 속에는 여러 가지 진동수 성분이 포함되어 있다. 이러한 성분을 추출, 정리하여 각 주파수 ω 에 대한 성분 강도의 분포 F(ω) 로써 재배열할 수 있다면 현상의 본질을 보다 쉽게 파악할 수 있다. 이처럼 주어진 f(t) 에서 F(ω) 를 구하는 조작이 푸리에 변환이며 수학적으로는 아래와 같이 표현된다[4].

$$F(\omega) = \frac{1}{\sqrt{2\,\pi}} \int_{-\infty}^{\infty} f(t) e^{-i\omega t} dt \tag{1}$$

푸리에 변환은 합성곱(Convolution)과 밀접한 관계를 가지고 있다. 합성곱 연산은 두 개의 함수 f 와 g 가 있을 때 이들의 합성곱을 수학 기호로는 f*g 으로 표시한다. 합성곱 연산은 두 함수 f 와 g 가운데 하나의 함수를 반전(reverse), 전이(shift) 시킨 다음, 다른 하나의 함수와 곱하고 이를 적분하는 것을 의미한다. 이를 수학 기호로 표시하면 다음과 같다[5].

$$(f*g)(t) = \int_{-\infty}^{\infty} f(\tau) g(t-\tau) d\tau \tag{2}$$

푸리에 변환과 합성곱 사이의 관계는, 두 개의 데이터에 대해 각각 푸리에 변환 연산을 실행한 다음 그 결과값들을 서로 곱해주고 이를 역 푸리에 변환하면 두 개의 데이터를 합성곱한 결과와 같으며, 이를 수식으로 나타내면 다음과 같다.

$$IFT\,[F(\omega) \times G(\omega)\,] = (f*g)(t) \tag{3}$$

여기서 F(ω) 와 G(ω) 는 각각 f 와 g 를 푸리에 변환한 결과이다. IFT(Inverse Fourier Transform) 푸리에 변환을 다시 역변환해주는 연산을 나타낸다. 푸리에 변환이 시간 영역에서 주파수 영역으로 신호를 분해하여 주파수 성분을 분석하는데 용이하다면 역 푸리에 변환은 반대로 주파수 영역을 시간 영

역으로 재배치하여 신호를 복구하는 성질을 가지고 있다. 즉 신호에 대한 주파수 및 위상 정보를 알면 기존의 신호를 정확하게 재구성할 수 있다.

2-3 웨이블릿 변환(Wavelet Transform)

신호는 어떠한 주파수와 위상을 가진 정현파들이 합쳐져서 만들어진 것이라고 할 수 있다. 푸리에 변환은 앞서 설명한 것과 같이 시간에 대한 함수 또는 신호를 주파수 성분으로 분해하는 연산을 말한다. 푸리에 변환은 연산 후 신호에 존재하는 주파수 성분들을 분석할 수 있는 장점이 있는 반면, 시간에 대한 정보가 없기에 언제 어느 시간에 그 해당 주파수가 존재하는지를 알 수 없다는 단점이 있다.

단시간 푸리에 변환(STFT: short-time fourier transform)은 위의 단점을 해결하기 위해 제안되었다. 시간에 따라 변하는 신호를 짧은 시간 단위로 분할한 다음 푸리에 변환 연산을 실행하는 것을 말한다. 결과적으로 짧은 시간 단위마다 어떠한 주파수 성분들이 존재하고 있는 지를 알 수 있다. 따라서 짧은 시간 단위 분할은 어떠한 시간 내에 어떠한 주파수가 존재하는지를, 긴 시간 단위 분할은 어떠한 주파수가 해당 시간 내에 존재하는지를 제공해준다. 즉 신호를 분할하는 시간 단위, 창(window) 넓이가 작으면 시간 분해능이 좋고 창의 넓이가 크면 주파수 분해능이 좋아진다.

그림 1은 STFT 에서 창의 넓이가 변화함에 따라 분해능이 어떤 차이를 가지는지를 보여주는 그림이다. 왼쪽 그림은 창의 넓이가 좁아짐에 따라 시간 분해능이 좋아지는 것을 보여주며 반대로 오른쪽 그림은 창의 넓이가 길어짐에 따라 주파수 분해능이 좋아지는 것을 보여주고 있다. 하지만 STFT 의 경우 시간과 주파수 분해능을 모두 만족시키는 적합한 지점을 찾기가 매우 어렵다. 이 때문에 새로운 연산 방법이 필요하게 되었고 그 결과로 웨이블릿 변환이 출현하게 되었다.

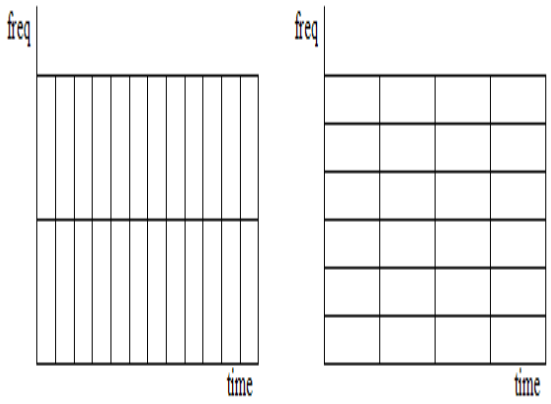

그림 1. 창(window)의 넓이에 따른 분해능 여부
Fig. 1. Resolution depending on the width of the window

그림 2는 STFT 와 웨이블릿 변환의 차이점을 보여주고 있다. 그림에서 알 수 있듯이 웨이블릿 변환은 STFT 와 달리 고주파 성분 신호에 대해서는 시간 해상도를 높이면서 주파수 해상도는 낮추는 반면에 저주파 신호에 대해서는 시간 해상도를 낮추면서 주파수 해상도를 높임으로써 시간 분해능 및 주파수 분해능 모두에서 강하다는 장점을 가지고 있다.

신호처리 전반에서 널리 응용되고 있는 웨이블릿 변환은 신호의 압축과 분석에 매우 유용하며[6], 웨이블릿의 다해상 분해(multiresoulution decomposition) 방법은 최근 분할(segmentation), 피치주기 검출 등의 다양한 방향으로 음성인식 시스템에 접목되어지고 있다[7].

웨이블릿 변환을 수식 기호로 표시하면 다음과 같다.

$$W_f(a,b) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{a}} \psi(\frac{t-b}{a}) f(t) dt \quad (4)$$

웨이블릿 변환의 기저 함수는 모 웨이블릿(mother wavelet) $\psi(t)$를 시간 축에서 b 만큼 변이(translation) 시키고, a 만큼 확대/축소(dilation) 시켜서 얻어낸다. a 가 1 보다 크면 기저 함수는 시간 축에서 좁아지며, a 가 1보다 작으면 넓어진다. 따라서 기저 함수들의 해상도를 조절할 수 있으며[8], 이러한 과정이 그림 2에 보여진다.

웨이블릿 변환에는 연속 웨이블릿 변환(CWT: Continuous Wavelet Transform)과 이산 웨이블릿 변환(DWT: Discrete Wavelet Transform)이 있으며, 일반적으로 DWT의 연산량이 더 적다. 실제 웨이블릿 시리즈는 DWT 수준에서 원신호를 복원하는데 필요한 만큼의 정보를 가지고 있기 때문에 많은 응용에서 DWT를 사용하고 있으며, 본 논문에서도 음성 데이터를 분석하는데 DWT를 사용하였다.

본 논문에서는 마이크를 이용하여 수집한 음성 데이터에 DWT 웨이블릿 변환을 적용하여 주파수 및 시간 대역에서 신호를 분석하였다. 그리고 신호의 고유한 특징을 나타내는 신호값 수십 가지를 추출한 뒤 이를 역변환 하여 음성 데이터의 특

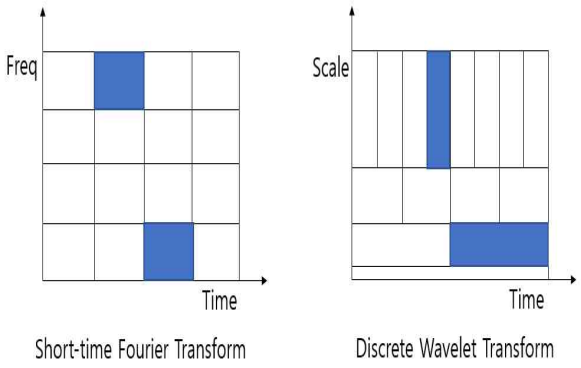

그림 2. STFT와 웨이블릿 변환 비교
Fig. 2. Comparison of STFT and Wavelet Transform

징 값들을 추출하였다. 이 특징 값들은 기계학습 ANN의 입력으로 사용되어 최적의 가중치를 얻는데 사용된다. 최적 가중치는 저장된 후에 암호화에서 사용된 음성 데이터와 복호화 과정에서 수집된 음성 데이터 간의 유사성 판별에 사용된다.

2-4 Machine Learning

기계학습은 경험적 데이터를 통해 학습하고 예측을 탐구하여 스스로 성능을 발전시키는 시스템과 이를 위한 알고리즘을 연구하고 구축하는 기술이다[9][10][11].

기계학습은 입력 데이터를 기반으로 출력 데이터를 예측하거나 결정을 이끌어낼 때 발생하는 오차값을 최소화 하도록 알고리즘을 구축하는 것이 매우 중요하다. 현재까지 이를 위한 많은 알고리즘이 만들어져 왔고 또한 만들어지고 있는 중이다. 본 논문에서는 기계학습의 기본 알고리즘 중 하나인 ANN(Artificial Neural Networks) 을 이용하였다. 따라서 ANN 및 이와 관련된 기술들에 대한 이해를 돕기 위해 아래에 간단한 설명을 추가하였다.

1) 신경망(Neural Network)

신경망은 인간이 뇌를 통해 문제를 처리하는 방법과 비슷한 방법으로 문제를 해결하기 위해 채택한 구조이다. 인간의 뇌에서 구조 조직인 뉴런(neuron)과 뉴런이 연결되어 일을 처리하는 것처럼, 수학적 모델로서의 뉴런이 상호 연결되어 네트워크를 형성할 때 이를 신경망이라 한다. 이를 생물학적인 신경망과 구별하여 인공 신경망(Artificial Neural Network)이라고도 한다. 신경망에서는 각 뉴런이 독립적으로 동작하는 처리기의 역할을 하기 때문에 병렬성(Parallelism)이 뛰어나고, 많은 연결선에 정보가 분산되어 있다. 따라서 몇몇 뉴런에 문제가 발생하더라도 전체 시스템에 큰 영향을 주지 않는 결함 허용(fault tolerance) 능력이 있으며, 주어진 환경에 대한 학습 능력이 있다. 이러한 특성은 인공 지능에 이용되고 있으며, 문자 인식, 화상 처리, 자연 언어 처리, 음성 인식 등 여러 분야에서 이용되고 있다[12].

그림 3은 ANN의 기본 구조를 보여준다. 입력층(Input layer)에 데이터가 입력되면 각각의 선에 주어진 가중치 값에 의하여 입력 중 어느 입력에 더 높은 비중을 둘 것인지가 결정된다. 이처럼 입력층에 입력된 각 입력은 각 가중치 값과 곱해진 후에 은닉층 (Hidden layer)에 연결되고 은닉층은 이제 입력 데이터가 되어 각각의 가중치를 가지고 다시 출력층(Output layer)에 연결 된다. 즉, 입력과 출력 사이가 가중치로 연결되어 입력 데이터 중 어느 데이터에 더욱 높은 비중을 둘 것인가가 결정된다. 그림 3에서 은닉층은 입력층의 출력층이라 할 수 있고 은닉층은 최종 출력층의 입력층이라 할 수 있다. 이 연산 관계를 수식으로 표현하면 다음과 같이 주어진다. (j는 Hidden 레이어에서 j 번째 순서의 노드를 나타내는 것이고 k 는 Input 레이어서 k

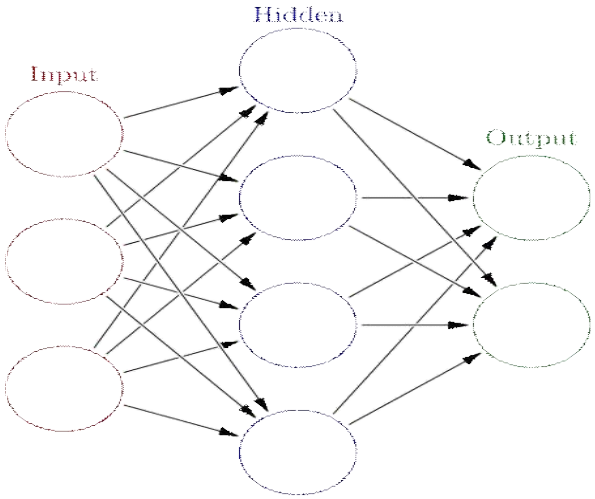

그림 3. 인공 신경망
Fig. 3. Artificial Neural Network

번째 순서의 노드를 나타내며, i는 Ouput 레이어에서 i 번째 순서의 노드를 나타낸다. H는 Hidden 레이어, X는 Input, Y는 Ouput, W는 가중치, b는 Bias이다.)

$$H_j = \sigma[\sum_{0}^{N} X_k W_{kj} + b] \quad (5)$$

$$Y_i = \sigma[\sum_{0}^{M} H_j W_{ji} + b] \quad (6)$$

먼저 각각의 층과 가중치가 곱해진 값을 전부 합치는 합 연산이 수행되며 합 연산 결과값이 얼마나 편향되어 있는지를 반영하기 위해 바이어스 값(b)을 마지막에 더해준다. 다음으로 이렇게 얻어진 결과값을 다음 층으로 넘길 것인지를 최종적으로 결정하기 위하여 활성함수(σ)를 적용한다. 활성함수 출력이 일정 수치를 넘지 못할 경우 결과값은 넘겨지지 않고, 일정 수치 이상인 경우에만 값이 넘겨진다. 활성함수로서는 시그모이드, ReLU, Softmax 등 다양한 함수들이 있으며 각각의 장단점을 고려하여 해당 알고리즘에 적합한 활성함수를 사용하여야 한다. 본 논문에서는 활성함수로서 Softmax 를 사용하였다. Softmax는 Multiclass classification 에 많이 쓰이는 함수이다.

최근 기계학습은 계층을 추가적으로 쌓는 것으로 더욱 높은 정확성을 얻어내는 심층 학습(Deep Learning)을 지향하는 반면, 계층이 깊어지면서 더욱 많은 메모리와 더욱 많은 시간을 할애하게 되었으나 본 논문에서는 간단한 ANN을 이용하여 심층 학습과 유사한 구조를 제안하였고 높은 정확성에 시간 할애와 메모리 사용을 단축하였다.

2) Softmax

Softmax 함수는 Multiclass classification 에 많이 쓰이는 활성함수로서 각각의 Class 에 대한 확률 밀도값을 계산하는 함수이다. 각 Class 에 대한 확률 밀도값을 모두 더하면 1이 되며, 가장 높은 확률 밀도값을 가지고 있는 Class 가 기대하던 Class 라고 분류하는 것이다. 위의 관계를 수식으로 표시하면 다음과 같다.

$$S(Y_i) = \frac{e^{Y_i}}{\sum_{0}^{K} e^{Y_i}} \quad (i = 0, 1, 2, ..., K)$$

(7)

최종 출력값을 지수함수에 넣은 뒤 분모에는 모든 Class 에 대한 값들을 모두 더한 값을 취하고 분자에는 각 Class 에 대한 값을 위치시킴으로써 해당 Class 의 확률 밀도를 구한다.

3) Cross Entropy

ANN 에 입력 데이터를 넣어 정확한 예측 값을 도출하기 위해서 사전에 필요한 전처리 과정이 있는데, 이는 기계학습의 메인이 되는 학습 단계이다. 학습 단계란 미리 준비해 둔 일련의 데이터들을 활용하여 판별 오차값을 최소화시키는 방향으로 가중치들을 수정해 나가는 단계를 말한다. 이 과정을 수식으로 나타내면 다음과 같다.

$$D(\widetilde{Y}, Y) = -\sum_{j} y_j \log(\widetilde{y_j})$$

(8)

위의 수식은 Softmax 를 활성함수로 사용하는 ANN 의 Cross Entropy 즉, 오차율을 나타낸다. 여기서 $\widetilde{Y}$ 는 Softmax 로 나타낸 확률 분포이고 Y 는 목표값이다. 본 논문에서는 판별을 용이하게 하기 위하여 참과 거짓으로 식별해주는 One - hot encoding 을 사용하였다. One - hot encoding 이란 목표값의 Class 가 0, 1, 2, 3, .., N 으로 표시될 때 오직 한 Class에만 참을 배정하고 나머지 Class 에는 거짓을 배정하는 것을 말하며 식별이 매우 용이해진 상태를 의미한다. 예를 들어 목표값의 Class 가 2개이며 이 목표값을 One - hot encoding 하게 되면 Class 0은 [1, 0]이 되고, Class 1은 [0, 1]이 된다. 만일 목표값이 [0, 1] 이라는 One - hot encoding 값으로 나타났다면 Softmax 의 값은 [0.1, 0.9] 내지 [0.4, 0.6] 과 유사한 값들이 나오게 될 것이다. 이제 Softmax 값이 [0.1, 0.9]이라는 가정하에 Cross Entropy 를 계산해보면 -0*log(0.1) -1*log(0.9) = -log(0.9) 가 되고 이 값은 0에 매우 가깝다는 것을 알 수 있다 (log(1) = 0). 즉 오차율이 거의 0 이라는 것을 의미한다. 식(5)와 식(6)에서 가중치가 갱신되면 Y_i 값이 변하고 이는 식(7)을 통하여 Softmax 출력값을 변동시킨다. 이는 다시 식(8)을 통하여 Cross Entropy 즉 오차율을 변동시킨다. 이처럼 학습 단계에서는 계산된 오차율을 감소시키는 방향으로 가중치를 계속해서 갱신해 나간다. 오차율을 줄이는 가중치 갱신 수식은 다음과 같다.

$$W_{\neq w} = W_{old} - \rho \frac{dE}{dW_{old}}$$

(9)

여기서 E 는 Cross Entropy 값이고 ρ 는 Learning rate 이다.

Ⅲ. 제안된 알고리즘 및 구현

제안된 암호화 및 복호화 알고리즘을 아주 간략히 설명하면 다음과 같다. 암호화 과정은 암호화 예정 데이터(입력 데이터)와 AV(Active Video) 장비로 획득한 Key 데이터를 각각 푸리에 변환한 후에 이들을 곱하여 진행된다. 한편 이의 복호화 과정은 먼저 동일인임을 확인한 후에 암호화 데이터를 푸리에 변환된 Key 데이터로 나누어 주고 그 결과를 역 푸리에 변환하여 진행된다.

Key 데이터로서 여러 가지 데이터들이 사용되어질 수 있지만 본 논문에서는 높은 보안성을 위하여 생체 신호에 가까운 안면 영상 데이터와 음성 데이터를 Key 데이터로 사용하였다. 그리고 복호화 과정 중 동일인 확인을 위하여 안면 영상 데이터 Key 간의 유사도 비교에서는 패턴매칭을 이용하였고 음성 데이터 Key 간의 유사도 비교에서는 웨이블릿 변환과 기계학습 방법을 이용하였다.

3-1 제안된 알고리즘

그림 4는 본 논문에서 제안된 알고리즘의 전체적인 구성도를 나타낸 것이다. 암호화 예정 데이터(입력 데이터) 및 안면 Key 데이터(Key A)를 준비하기 위하여 2대의 Webcam 장비를, 그리고 음성 Key 데이터(Key B)를 위해서는 1대의 마이크를 사용하였다. 암호화 및 복호화 전체 과정에서의 데이터 사전 작업 및 신호 처리를 위하여 LabVIEW를 사용하였다.

그림 4의 암호화(Encrypt) 과정에서는 암호화 예정 데이터, 안면 데이터, 그리고 음성 데이터 각각을 푸리에 변환한 후에 이들을 차례로 곱하여 암호화 결과물인 출력 값 Y(ω) 를 생성한다. 한편 복호화(Decrypt) 과정에서는 유사도 판별을 통하여 동일인임이 확인된 경우에 한하여 곱해주었던 Key 들을 다시 나누어 주고 최종적으로 역 푸리에 변환을 통하여 입력 데이터를 다시 복원한다. 위에서 유사도 판별은 복호화 초기에 수집된 Key C 와 Key D 데이터와 암호화 과정에서 사용되었던 Key A 와 Key B 데이터 간의 유사성을 조사하는 것이다.

Key A 와 Key C가 일정 수준 이상의 유사성을 가지며 Key B 와 Key D 도 일정 수준 이상의 유사성이 확인(Key A ≒ Key C , Key B ≒ Key D)되면 즉 동일인임이 확인되면 수집된 Key

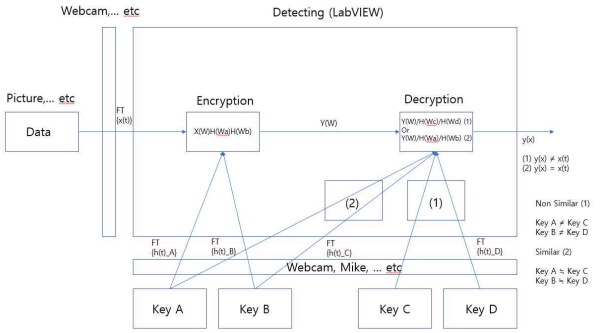

그림 4. 전체적인 구성도
Fig. 4. Overall block diagram

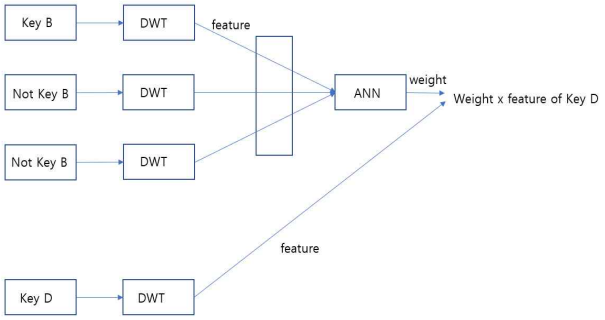

그림 5. 음성 데이터 Key에 대한 유사도 판별 구성도
Fig. 5. Block diagram for Similarity Determination in case of speech data Key

C 와 Key D 를 사용하는 대신에 암호화 과정에서 사용되었던 Key A 와 Key B 를 다시 사용하여 복호화를 성공적으로 수행할 수 있다. 제안된 방식은 암호화에서 사용된 2개의 생체 신호 Key 들과 일정 수준 이상으로 유사성을 가진 2개의 생체 신호 Key 가 복호화 과정에서 확보되어야만 원래의 데이터를 복원할 수 있다는 점에서 1개의 정적인 Key를 사용하는 이전 방식들 보다 더 높은 보안성을 기대할 수 있다.

본 논문에서는 Key A 와 Key C 로서 사용된 안면 영상 데이터 간의 유사도 판별을 위하여 패턴 매칭 방법을 사용하였다. 한편 Key B 와 Key D 로서 사용된 음성 데이터 간의 유사도 판별과정은 그림 5에 나타내었다. 먼저 암호화 과정에서 사용된 Key B 와 비슷한 데이터들(즉 동일한 단어에 대하여 여러 번 녹음한 음성 데이터들)을 준비한다. 그리고 Key B 와 완전 다른 음성 데이터들(즉 다른 단어들에 대하여 녹음한 음성 데이터들)도 준비한다. 이는 인공지능 ANN 학습을 위함이다. 다음으로 준비된 각각의 음성 데이터에 대한 웨이블릿 변환을 통하여 음성 데이터별 특징 값들을 추출하고 이들을 ANN 의 입력 데이터로 활용한다. 이어서 ANN 을 이용한 학습을 통하여 최적의 가중치를 확보하고 이를 저장한다. 이제 복호화 과정에서 마이크를 통하여 실시간으로 얻어진 음성 데이터 Key D 에 대한 웨이블릿 변환을 통하여 특징 값을 추출하고 이를 저장된 최적의 가중치와 곱 연산을 수행한다. 이 결과 값과 Key B에 대하여 동일한 과정을 수행하여 얻어진 값을 비교하여 Key B와 Key D 간의 유사도를 판별한다.

그림 5에서 ANN을 통한 가중치 추출 과정은 Key B 에 대한 푸리에 변환 전에 실행된다. 따라서 그림 4에서 마이크를 이용하여 Key B 를 수집한 직후에 진행해주어야 한다.

3-2 암호화(Encryption)

암호화는 입력 데이터와 안면 데이터(Key A) 및 음성 데이터(Key B)를 각각 푸리에 변환한 후에 이들을 순서대로 곱하여 진행된다.

입력 데이터 및 Key A는 2차원 데이터 값으로서 480 x 640 pixel 사이즈를 채택하였고 푸리에 변환 한 후 동일한 좌표값에 해당하는 값들을 서로 곱해주었다.

음성 데이터(Key B) 의 경우 시간에 따라 변하는 1차원 데이터로 구성된다. 그런데 암호화를 위해서는 1차원 데이터를 2차원으로 변경해야 하므로 길게 나열된 1차원 배열을 분할하여 다음 행에 쌓아두는 방법으로 2차원 변경을 실행하였다. 음성 데이터(Key B) 의 샘플링 속도를 16000Sample/sec(나이키스트 샘플링 기본 속도보다 10배)로 하여 획득한 1차원 배열 데이터(1x16000)를 480개씩 잘라서 순서대로 다음 행에 쌓아두는 방식을 채택하였다. 부족한 부분은 다시 음성 데이터의 16000개 데이터 값을 불러와 차례로 채워주었다. 이렇게 만들어진 480 x 640의 음성 데이터(Key B) 를 푸리에 변환하였고 이를 입력 데이터와 Key A 를 각각 푸리에 변환한 후 곱해준 결과에 다시 곱해주어 최종적으로 암호화된 결과값 Y(ω) 를 얻었다.

입력 데이터와 Key A 의 경우 Webcam을 이용하여 영상 데이터를 확보하였고 Key B 의 경우 마이크를 이용하여 수집하였다. 다만 암호화 단계에서는 실시간 데이터를 사용한 것이 아니라 미리 수집된 데이터들을 불러 와 사용하였다.

앞에서 언급하였듯이 암호화 과정에서는 Key B 를 포함한 수 많은 음성 데이터에 웨이블릿 변환을 적용하여 각각의 특징 값들을 추출하고 이들을 테스트 데이터로 사용한 ANN 학습을 진행하여 최적의 가중치를 확보한다. 그리고 복호화 과정에서 이를 다시 사용하기 위하여 저장해둔다.

암호화 과정에서 ANN 학습을 위해 사용된 데이터의 Class 는 3 개이다. 즉, 묵음(No sound) 데이터, 상승(Up) 음 데이터, 하강(Down) 음 데이터이며 각 Class 당 10개씩 총 30개의 데이터를 사용하였다. 본 논문에서는 하강 음 데이터를 Key B 로 사용하였다. 학습 횟수는 약 8500만 번이었으며, 약 98% 의 정확도와 평균적으로 0.175 미만의 Cross Entropy 값이 얻어졌다. 본 논문에서는 Cross Entropy 값이 0.3 이상일 때 판단 오차가 출력되도록 하였다.

위 수치들은 그림 6과 그림 7에서 쉽게 확인될 수 있다. 그림 6은 학습을 위해 사용되었던 특징 값들을 보여주고 있다. 첫 번째 테두리의 10개의 행은 10개의 묵음 데이터, 두 번째 테두리의 10개의 행은 10개의 상승 음 데이터, 세 번째 테두리의 10개

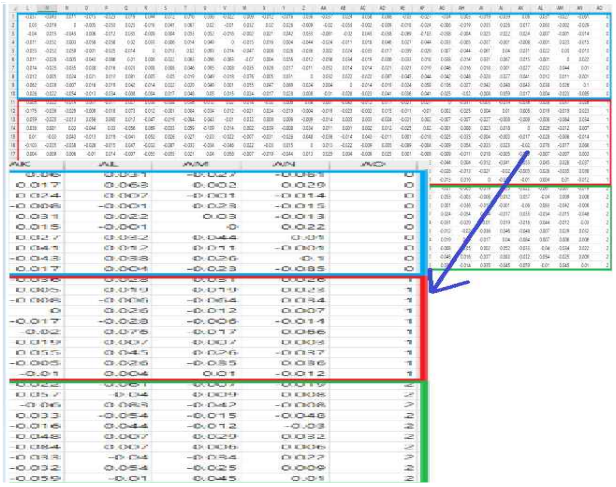

그림 6. 학습 데이터
Fig. 6. Learning Data

의 행은 10개의 하강 음 데이터에 해당하며 각 행의 40개 데이터는 특징 값들을 나타낸다. 본 논문에서 암호화에 사용된 음성 데이터(Key B) 는 하강 음을 녹음한 것으로 녹색 테두리에 있는 각 행의 40개 데이터는 특징 값들이다. 복호화 때 실시간으로 수집된 음성 데이터(Key D) 가 하강 음이 아닌 경우에 이를 정확하게 판단할 수 있도록 하기 위하여 하강 음이 아닌 데이터들이 필요하며 나머지 20개 행의 데이터가 이에 해당한다. 그림 6의 액셀표의 가장 우측에 보이는 0, 1, 2는 각각 묵음, 상승 음, 하강 음 데이터의 Class 를 나타내며, 해당 Class 를 One - hot encoding 하여 학습을 진행하였다.

그림 7은 LabVIEW 프로그램을 사용하여 수행한 학습의 결과를 보여주고 있다. 그림은 마지막 학습 출력값을 보여주고 있는데, One - hot encoding 과정을 사용하므로 Class 는 이진 데이터 값으로 표시되며 Target 부분에 이 결과값이 보여진다. Target 부분은 [0, 0, 1] 로 나타나 있고 이것의 오차율을 표현하는 Cross Entropy를 최소화 하기 위해서는 Softmax 활성함수에서 출력된 확률 분포값이 세 번째 Class에서 가장 높게 나와야 한다는 것을 알 수 있다. 그림에서 Softmax 출력 값을 확인해보

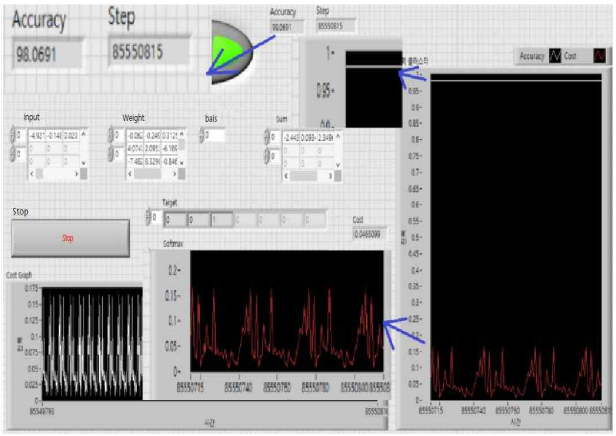

그림 7. 기계학습 결과
Fig. 7. Machine Learning Result

면 예상대로 세 번째 Class에서 가장 높은 확률 분포값(약 0.898)이 나타난 것을 알 수 있다.

그림 7을 보면 사용된 입력 데이터, 추출된 최적 가중치 등 여러 값들을 확인할 수 있다. 우측의 출력 클러스터 그래프를 보면 정확도와 오류값이 나타나 있는데, 정확도는 1에 근접한 것을 확인할 수 있고 (약 98%), 오류값은 모두 약 0.175 미만인 것을 확인할 수 있다.

3-3 복호화(Decryption)

복호화 과정에서는 Webcam 과 마이크를 사용하여 실시간으로 Key C와 Key D 를 수집하였고 동일인임을 확인하기 위하여 암호화 단계에서 얻어진 Key A 및 Key B 와의 유사성을 확인하였다. 그리고 동일인 판정 여부에 따라 기존의 데이터인 Key A 및 Key B 를 다시 사용하여 원래 데이터를 복원할 것인지(복원 성공) 아니면 Key C 및 Key D 를 그대로 사용할 것인지(복원 실패)를 결정하게 된다.

복호화 과정에서 다시 수집된 안면 데이터(Key C) 의 사이즈는 암호화 과정과 동일하게 480 x 640 pixel 으로 채택하였다. 복호화에서는 Webcam 을 이용하여 실시간으로 안면을 계속 촬영하고 있기 때문에 수집되는 데이터는 엄밀히 따지면 동영상 데이터이다. Key A 와 Key C 간의 유사성을 판별하기 위하여 LabVIEW 프로그램에 내장된 Vision 이라는 함수를 사용하였다. 이는 ROI(Region of Interest) 지정, 패턴 매칭, 그리고 거리 측정 등 다양한 기능을 제공하고 있다. 본 논문에서는 Vision 함수로 Key A의 ROI를 지정하여 이 영역 안의 pixel 데이터를 저장해두고 실시간으로 수집되는 동영상 데이터(Key C)의 pixel 값과 비교하였다. 이때 유사성의 한계 수치로서 0.9의 확률 값을 채택하였고, 3초 동안 0.9 이상의 확률 값이 지속되면 Key A ≒ Key C 으로 판정하였다. 그리고 Key C 를 대신하여 암호화 때 사용되었던 Key A 를 복원에 다시 사용하였다. 만약 3초 이내에서 0.9 미만의 유사성이 확인되면 시간은 다시 0초로 리셋되고 이때부터 다시 3초 동안 유사성 판별이 시작된다.

복호화 때 마이크를 통해 실시간으로 얻어지는 음성 데이터(Key D) 의 경우도 동일인 확인을 위해 Key B 와의 유사성 판별이 필요하다. 본 논문에서는 이를 위하여 웨이블릿 변환과 기계학습의 ANN 및 활성함수 Softmax 를 사용하였다. 음성 데이터(Key D) 에 웨이블릿 변환을 적용하여 주파수 및 시간 대역에서의 신호를 분석하였고 그 중에서 Key D의 특징을 잘 나타내는 신호 값들을 특징 값으로 추출하였다. 이 특징 값들을 암호화 시에 저장해 두었던 최적 가중치 값들과 곱하여 목표값이 제대로 나오는 지를 확인하는 것이 유사성을 판별하는 방법이다. 유사성을 위한 목표값 한계치로서는 0.9를 설정하였다.

Ⅳ. 실험 결과

실험 결과는 암호화 및 복호화에 성공한 경우와 실패한 경우로 나뉠 수 있다. 본 장에서는 두 경우에 대하여 모두 살펴보고자 한다. 그런데 실패하는 경우는 다시 세 가지 경우로 나뉠 수 있다. 첫 번째는 안면 데이터(Key C) 는 유사성이 확인되었는데, 음성 데이터(Key D) 가 유사하지 않은 것으로 판별된 경우이다. 두 번째는 음성 데이터(Key D) 는 유사한데, 안면 데이터(Key C) 가 유사하지 않은 경우이다. 세 번째는 안면 데이터(Key C) 와 음성 데이터(Key D) 모두 유사하지 않은 것으로 판별된 경우이다. 이 세 가지 중 어느 하나가 발생할 경우 복호화에 실패하여 입력 데이터와 상이한 결과가 초래된다. 본 논문에서는 세 가지 경우 중 한 가지 경우에 대하여만 살펴보도록 하겠다.

그림 8은 복호화에 성공한 경우를 , 그리고 그림 9는 복호화에 실패한 경우를 보여준다. 이 그림들은 LabVIEW 프로그램으로 설계하여 만든 실시간 모니터링 창이다. 그림 8과 그림 9의 상단에는 입력 데이터(Input)로서 National Instrument 로고 사진 데이터를, 안면 데이터(Key A) 로서는 Pattern Key 에서 보여진 얼굴 사진을, 그리고 음성 데이터(Key B) 로서는 Speech Key에서 보여진 Class 2의 하강 음 데이터를 나타내었다. 또한 암호화 결과를 Encryption 에 나타내었다. 그림 8과 그림 9의 하단에서는 복호화를 위해 수집된 안면 데이터(Key C)로서 Pattern Compare Key에서 보여진 얼굴 사진을, 음성 데이터(Key D)로서 Speech Compare Key에서 보여진 음성 데이터를, 그리고 복호화된 결과는 Decryption에 각각 나타내었다. 앞 장에서 설명한 것처럼 Encryption은 Input과 Pattern Key 및 Speech Key 를 각각 푸리에 변환한 뒤 차례로 곱 연산을 수행하여 얻어진다. 복호화(Decryption)를 위해서는 먼저 동일인 여부 판정을 위하여 Pattern Key(Key A)와 Pattern Compare Key(Key C), 그리고 Speech Key(Key B)와 Speech Compare Key(Key D) 간의 유사성을 판별한다. 만일 유사성이 모두 인정되면 암호화 결과(Encryption)를 다시 Pattern Key(Key A)와 Speech Key(Key B) 의 푸리에 변환 값으로 각각 나누어 주고 이 결과의 역 푸리에 변환을 통하여 복원(Decryption)이 완료된다. 그림 8에서 암호화에 사용된 안면 데이터(Key A)와 복호화에서 실시간으로 얻어진 안면 데이터(Key C)가 유사성 판별을 통과하고, 암호화에 사용된 Class 2 하강 음 데이터(Key B)와 복호화에서 실시간으로 얻어진 음성 데이터(Key D)가 유사성을 통과할 때 하단 Decryption에 National Instrument 로고 사진이 잘 복원되었음을 알 수 있다.

그림 9에서는 복호화에 실패한 경우를 나타내었다. 암호화에서 사용된 안면 데이터(Key A) 와 복호화에서 사용된 실시간 안면 데이터(Key C) 간의 유사성 판별은 통과되었지만, 암호화에서 사용된 Class 2 하강 음 데이터와 복호화에서 수집된 Class 0 묵음 데이터 간에 유사성이 없으므로 복호화에 실패하여 Decryption 에서는 National Instrument 로고 사진이 아닌 엉뚱한 데이터가 나타남을 확인할 수 있다.

본 논문에서 제안된 알고리즘은 LabVIEW 프로그램을 사용하여 설계되었기 때문에 LabVIEW 의 강한 장점 중 하나인 실시간 모니터링이 가능하며 암호화 및 복호화에 걸리는 시간을 0.001 초 단위로 조정할 수 있다. 실험에서 복호화에 소요되는 시간은 안면 데이터의 유사성 판별에 소요되는 시간(본 실험에서는 3초로 설정하였지만 단축 조정이 가능함)과 음성 데이터를 녹음하고 샘플링하는데 걸리는 시간(약 1~2초)을 감안하더라도 약 5초 이내인 점을 고려해 볼 때 제안된 알고리즘의 실용화에는 큰 문제가 없는 것으로 사료된다.

본 논문에서 제안된 암호화 및 복호화 알고리즘은 안면 및 음성 데이터와 같은 이중 생체 데이터를 Key로 사용하므로 기존의 알고리즘보다 보안성을 더욱 높일 수 있으며, 실시간으로 입력되는 동적인 데이터를 Key로 사용한다는 점에서 더욱 편리함을 제공할 수 있다. 나아가 제안된 알고리즘은 푸리에 변환을 이용하므로 Key 로 사용될 수 있는 데이터가 본 실험처럼 영상 및 음성으로 제한되지 않고 다양하게 선택될 수 있다는 범용성도 보여준다.

본 논문에서는 웨이블릿 변환을 이용한 특징 값 추출과 기

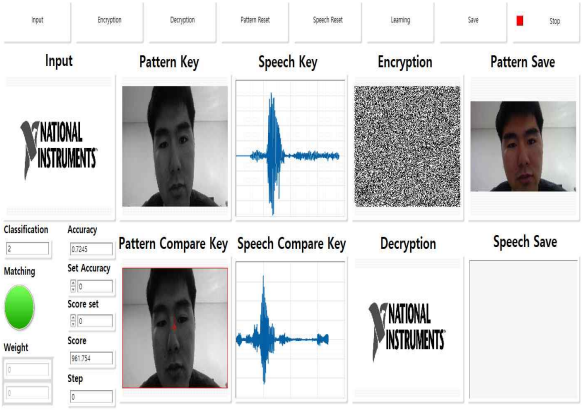

그림 8. 복호화 성공
Fig. 8. Encryption & Decryption Success

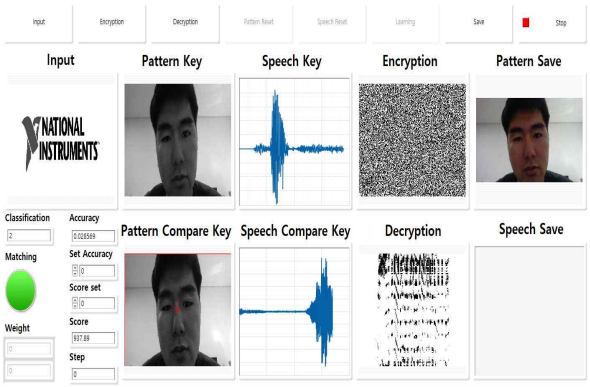

그림 9. 복호화 실패
Fig. 9. Encryption & Decryption Failed

계학습의 ANN 알고리즘을 함께 사용하여 복호화 과정에서 음성 데이터를 높은 정확도로 인식할 수 있다는 것을 보였다. 그러나 기계학습의 특성상 학습하지 않은 데이터가 Key 로 입력되어지면 종종 오류를 범한다는 단점을 확인할 수 있었다. 따라서 새로 사용되는 Key 에 대한 학습이 계속 진행되어야 하는 것이 과제로 남겨졌다. 또한 음성 데이터로부터 정확한 특징 값 추출을 위하여 잡음제거 등과 같은 전처리 과정에 보다 우수한 장비가 사용될 필요성도 제기되었다.

Ⅴ. 결 론

본 논문에서는 생체 데이터인 안면 데이터와 음성 데이터를 Key로 사용하여 이중으로 암호화 및 복호화를 수행하는 알고리즘을 제안하였다. 그리고 제안된 알고리즘을 LabVIEW 프로그램을 사용하여 실시간으로 구현하였다. 암호화는 입력 데이터와 Key 데이터들을 각각 푸리에 변환한 뒤 차례로 곱 연산을 수행하여 진행된다. 복호화는 두 단계로 진행되는데 처음에는 동일인 확인을 위하여 새로 수집된 Key 데이터들과 암호화에서 사용된 Key 데이터들간의 유사성을 판별한다. 이 과정에서 웨이블릿 변환을 사용하여 음성 Key 데이터의 특징 값을 추출하고 여기에 기계학습 ANN을 적용하여 판별에 대한 정확성을 높였다. 다음으로 유사성이 모두 인정될 때 암호화 결과값을 암호화에서 사용된 각 Key의 푸리에 변환 값으로 다시 나누어 주므로써 입력 데이터를 복원한다. 실험에서 복호화에 소요되는 시간은 안면 데이터의 유사성 판별에 소요되는 시간(본 실험에서는 3초로 설정)과 음성 데이터를 녹음하고 샘플링하는데 걸리는 시간(1~2초)을 합하여 약 5초 이내이었다. 이로부터 제안된 알고리즘의 실용화에는 큰 문제가 없는 것으로 사료된다. 그리고 실시간으로 입력되는 안면과 음성 생체 데이터를 모두 Key로 사용한다는 점에서 기존의 알고리즘보다 더 높은 보안성과 편리함을 제공할 수 있다. 나아가 제안된 알고리즘은 푸리에 변환을 이용하므로 Key 로 사용될 수 있는 영역이 다양하여 범용성이 풍부하다. 또한 ANN은 최근 계층을 깊게 쌓는 심층 학습으로 진화하면서 영상 및 음성인식이나 번역 등의 다양한 분야에서 뛰어난 성능을 보여주고 있다[13]고 알려져 있는 반면, 계층이 깊어지는 것에 따라 학습을 마무리 하는 데 시간이 오래걸리고 장비의 성능이 더욱 높아져야 하는데 본 논문에서 제안한 방법으로는 기존의 ANN을 이용하여 특징 값을 뽑아내어 계층이 깊게 쌓여있는 딥 러닝과 유사한 구조를 만들어내었다는 것에서 시간 단축과 적은 메모리 사용으로도 높은 정확성을 얻어냈다.

감사의 글

본 연구는 2017년도 강원대학교 대학회계 학술 연구조성비(관리번호-520170071)에 의하여 이루어진 연구로서, 관계부처에 감사드립니다.

본 연구는 과학기술정보통신부 및 정보통신기획평가원의 대학ICT연구센터지원사업의 연구결과로 수행되었음.(IITP-2019-2018-0-01433)

참고문헌

[1]D.K.Lee “Encryption and decryption using Fourier transform”, IKEEE, Hongik University, pp. 133, 2019. 08.

[2]M.S.Lee “Modern Cryptology”, kyowoo, ‘99.

[3]S.J.Choi “Graphical Programming LabVIEW introduction”, pp. 14, 1998. 12.

[4]Fourier transform, Electrical Term Dictionary, https://terms.naver.com/entry.nhn?docId=597869&cid=42340&categoryId=42340

[5]Wikipedia ,convolution, https://ko.wikipedia.org/wiki/%ED%95%A9%EC%84%B1%EA%B3%B1#%EC%9D%B4%EC%82%B0_%ED%95%A9%EC%84%B1%EA%B3%B1

[6]G.S.Seo, "(An) enhanced text-prompt speaker recognition using DTW," Master's Thesis, Chonbuk National University 1999. 02.

[7]S.M.Kim, G.S.Seo, C.K.Kim “New Speech Feature Extraction using DWT based Nonuniform filter banks,” National IT Industry Promotion Agency, pp. 1, 2000.

[8]G.T.Kim, "Design of A WideBand Speech Codec Using Wavelet Transform," Master's Thesis, Inha University, pp. 7, 2000. 02.

[9]B.G.Lee, J.T.Lim, J.S.You, “Utilization of Social Media Analysis using Big Data,” The Journal of the Korea Contents Association, Vol. 13, No. 2, pp. 211-219, 2013.

[10]D.K.Choi, J.O.Park, "The Application Method of Machine Learning for Analyzing User Transaction Tendency in Big Data environments," Journal of the Korea Institute of Information and Communication Engineering, Vol. 19, No. 10, pp. 2232-2240, Oct. 2015.

[11]Maching learning, WIKIPEDIA, Available : https://en.wikipedia.org/wiki/Machine_learning/, Oct. 2018.

[12]neural network, Computer Internet IT Terms Dictionary, https://terms.naver.com/entry.nhn?docId=830577&cid=42344&categoryId=42344

[13]H.J.Lee “A Study on the Uncertainty of Rules Extracted from Artificial Neural Networks,” Master's Thesis, Korea University, pp. 1, 2018.12.

이덕규(Duk-Kyu Lee)

2011년～2017년: 강원대학교 전자공학과 (공학사)
2017년～현　재: 강원대학교 전자공학과 (공학석박사통합과정)

※관심분야 : 광통신, 기계학습 등

엄진섭(Jinseob Eom)

1978년～1982년: 서울대학교 전자공학과 (공학사)
1982년～1984년: 서울대학교 전자공학과 (공학석사)
1987년～1990년: 미국 Texas A&M Univ. Dept. of Electrical Eng. (공학박사)

1991년～1992년: 한국통신 선임연구원
1992년～현　재: 강원대학교 전자공학과 교수
※관심분야 : 광통신, 광센싱, 생체의료광학 등