

FP-ICA에 의한 잡음에 강건한 디지털영상 워터마킹

조용현, 홍성준*
대구가톨릭대학교 컴퓨터정보통신공학부
e-mail:yhcho@cataegu.ac.kr

Robust Watermarking in Noise for Digital Image by Using FP-ICA

Yong-Hyun Cho and Seong-Jun Hong*
School of Computer and Information Communications. Eng.,
Catholic Univ. of Daegu

요 약

본 논문에서는 고정점 알고리즘 독립성분분석을 이용하여 잡음에 강건한 디지털영상의 워터마킹을 제안하였다. 여기서 고정점 알고리즘은 워터마킹의 추출과정에서 빠른 추출과 잡음에 강건한 개선된 추출성능을 얻기 위함이고, 독립성분분석의 이용은 추출과정에서 워터마크의 위치나 크기, 원본과 키 영상 등에 대한 사전 지식의 요구를 없애기 위함이다. 제안된 기법을 256x256 픽셀의 레나 원 영상, 키 영상, 그리고 문자 워터마크에 적용한 결과, 잡음과 같은 공격에 강하면서도 워터마크의 검출 및 추출과정에 원본 영상들에 대한 사전지식이 요구되지 않았다.

1. 서론

디지털 미디어의 빠른 보급에 따라 디지털화된 콘텐츠의 보급 또한 용이해지고, 이들의 불법적인 복사나 수정, 그리고 전송 등이 문제로 대두되고 있다 [1-4]. 디지털 워터마킹(watermarking)은 일반적으로 저작권 보호나 복사자의 추적, 그리고 내용의 조작이나 변경 등을 확인하는데 널리 이용되고 있다 [1,2]. 일반적으로 워터마킹을 위한 선행조건으로 워터마크는 투명하게 삽입되어 시각적이나 청각적으로 인지될 수 없는 비가시성과 여러 가지 영상처리에도 워터마크 정보는 그대로 유지되는 견고성이 요구된다. 또한 영상의 저작자나 사용자가 유일하게 특정되어야 하는 유일성과 정확한 키(key)를 알아야 워터마크의 검출과 추출이 가능한 보안성 등의 조건들도 만족되어야 한다.

일반적으로 워터마킹은 생성된 워터마크와 키를 삽입하는 과정과 역으로 검출 및 추출하는 과정으로 구성된다. 삽입과정은 원 영상에 워터마크를 삽입하는 과정이며, 검출과정은 워터마크의 존재여부를 확인하는 과정이고, 추출과정은 삽입된 워터마크를 원래의 워터마크로 재구성하는 과정이다. 특히 검출 및 추출 과정에는 워터마크의 위치나 크기 등에 대한 사전 지식이 요구되며, 원 영상과 키에 대한 지

식도 요구되고 있다. 따라서 실시간으로 워터마킹을 수행하기 위해서는 이들에 대한 사전 지식이 요구되지 않아야 한다.

한편 독립성분분석(independent component analysis : ICA)은 신호의 통계적 성질을 고려한 다차원의 선형변환 기법이다[6-9]. 이는 통신이나 신호 및 영상 처리 등에서 은닉신호분리(blind source separation : BSS)나 특징추출에 널리 이용되고 있다. 특히 BSS는 혼합된 신호로부터 원래의 신호를 추출하는 것으로 다양한 분야에 응용될 수 있다.

본 연구에서는 뉴우턴(Newton)법[10]에 기초한 고정점(fixed-point : FP) 알고리즘의 ICA(FP-ICA)를 디지털 영상의 워터마킹에 이용하는 방안을 제시한다. 여기서 FP 알고리즘의 이용은 빠른 추출과 함께 잡음에 강건한 우수한 추출성능을 얻기 위함이고, ICA의 이용은 워터마크의 검출/추출 과정에 워터마크의 위치나 크기, 원 영상과 키 등에 대한 사전지식의 요구를 없애기 위함이다. 제안된 기법을 256x256 픽셀(pixel)의 원본 영상을 보호하기 위한 워터마킹에 적용하여 그 타당성과 잡음공격에 따른 성능을 확인하였다.

2. 뉴우턴법의 FP-ICA에 의한 워터마킹

은닉신호의 분리나 특징추출에 널리 이용되고 있는 ICA는 m 개의 입력신호 \mathbf{s} 로부터 선형적으로 혼합된 n 개의 혼합신호 \mathbf{x} 가 알려져 있을 때, 혼합된 신호로부터 역으로 m 개의 독립인 입력신호를 찾는 기법이다. 혼합신호와 입력신호와의 관계는

$$\mathbf{x} = \mathbf{A}\mathbf{s} + \mathbf{n} = \sum_{i=1}^m s(i)\mathbf{a}(i) + \tilde{\mathbf{N}} \quad (1)$$

로 정의된다. 여기서 잡음 $\tilde{\mathbf{N}}$ 은 보통 입력신호와 구별되지 않기 때문에 생략할 수도 있으며, 혼합행렬 $\mathbf{A}=[\mathbf{a}(1), \mathbf{a}(2), \dots, \mathbf{a}(m)]$ 으로 $\mathbf{a}(i)$ 는 ICA의 basis vector이다. 결국 ICA는 알려진 혼합신호로부터 혼합행렬의 역행렬 $\mathbf{A}^{-1}(=\mathbf{W})$ 을 찾는 기법이다.

일반적으로 좀 더 효과적인 ICA를 위해 영평균과 whitening의 전처리 과정을 수행한다[6-8]. 신호의 영평균은 신호벡터 \mathbf{x} 에서 평균값 \mathbf{x}^* 를 뺀 차로

$$\mathbf{x} = \mathbf{x} - \mathbf{x}^* \quad (2)$$

와 같다. whitening은 신호벡터 \mathbf{x} 의 공분산행렬이 단위행렬값을 갖도록 함으로써 구할 수 있다. 즉, $E\{\mathbf{x}\mathbf{x}^T\}=\mathbf{I}$ 가 되도록 한다. 이를 위해 PCA가 널리 이용되며, 이때의 whitening 행렬 \mathbf{V} 는

$$\mathbf{V} = \mathbf{D}^{(-1/2)}\mathbf{E}^T \quad (3)$$

이다. 여기서 \mathbf{D} 는 고유값으로 구성되는 대각행렬이며, \mathbf{E} 는 고유벡터로 구성되는 직교행렬이다.

독립성분분석을 위한 FP 알고리즘은 엔트로피 최적화에 기초한 알려진 기법 중 가장 빠른 학습속도를 가진다[6-9]. 이는 신호 내에 포함된 상호정보를 최소화 함으로써 ICA의 해를 구하는 기법이다. 특히 알고리즘의 한 단계에 많은 수의 데이터들이 이용되는 batch mode로 계산을 수행하며, 신경망으로 구현될 수 있는 알고리즘이다. 이러한 FP 알고리즘을 위한 역혼합행렬 \mathbf{W} 경신에는 기존의 수치적 계산법인 뉴우턴법 등이 이용될 수 있다. 뉴우턴법의 FP 알고리즘에 의한 근사화된 반복기법의 역혼합행렬 \mathbf{W} 경신식은

$$\begin{aligned} \mathbf{W}^+ &= \mathbf{W} - [E\{\mathbf{x}\mathbf{g}(\mathbf{W}^T\mathbf{x})\} - \beta\mathbf{W}] / [E\{\mathbf{g}'(\mathbf{W}^T\mathbf{x})\} - \beta], \\ \mathbf{W}^* &= \mathbf{W}^+ / \|\mathbf{W}^+\| \end{aligned} \quad (4)$$

와 같으며, \mathbf{W}^* 는 \mathbf{W} 의 새롭게 경신된 값이고, $\beta = E\{\mathbf{W}^T\mathbf{x}\mathbf{g}(\mathbf{W}^T\mathbf{x})\}$ 이다. 식 (4)의 경신식에서 $\mathbf{g}(\cdot)$ 는 비선형함수로 $(\cdot)^3$ 과 $\tanh(\cdot)$ 의 함수를 이용한다. 여기

서는 $(\cdot)^3$ 를 이용하였다.

이상의 ICA가 가지는 은닉성분분석의 속성을 이용하면 워터마크된 영상으로부터 워터마크와 원 영상, 그리고 키 영상의 분리가 가능하다. 결국 워터마크된 영상들을 입력으로 ICA를 수행함으로써 은닉되어 있는 원 영상, 워터마크, 그리고 키 영상들을 각각 추출할 수 있다. 이때 기존의 워터마크 검출 및 추출에서와는 달리 원 영상이나 키 영상, 워터마크 등의 사전지식을 요구하지 않는다. 본 논문에서는 먼저 워터마크 삽입과정으로 식 (5)와 같이 선형 결합된 영상을 생성한다.

$$\mathbf{x} = \mathbf{s} + \mathbf{a}\mathbf{K} + \mathbf{b}\mathbf{M} \quad (5)$$

여기서 \mathbf{s} , \mathbf{K} , \mathbf{M} 은 각각 원 영상, 키 영상, 워터마크이고, \mathbf{a} 와 \mathbf{b} 는 각각 작은 가중치들이다. 이때 워터마크된 영상에서 삽입된 키 영상과 워터마크가 보이지 않도록 하기 위해서 원 영상의 에너지보다 충분히 작은(보통 10 - 100배 정도) 에너지의 키 영상과 워터마크를 이용한다. 또한 ICA에서의 입력신호 수가 원래 신호의 수와 같거나 그 이상이어야 하는 조건에 따라 다음의 식 (6)과 같이 3개의 입력영상을 생성한다.

$$\mathbf{x}_1 = \mathbf{x}, \quad \mathbf{x}_2 = \mathbf{x} + \mathbf{c}\mathbf{K}, \quad \mathbf{x}_3 = \mathbf{d}\mathbf{s} \quad (6)$$

여기서 \mathbf{c} 와 \mathbf{d} 는 임의의 작은 실수값이다.

따라서 FP-ICA를 이용한 제안된 워터마킹의 수행과정을 요약하면 다음과 같다.

단계 1 : 워터마크 및 키 영상을 생성한다.

단계 2 : 선형결합에 의해 원 영상에 워터마크와 키 영상을 삽입한다.

단계 3 : 워터마크된 영상으로부터 식 (6)에 의해 3개의 ICA 입력신호를 생성한다.

단계 4 : 고정점 알고리즘 ICA를 수행한다.

(1) $\mathbf{W}_{old} = \mathbf{W}$ 를 계산한다.

(2) $\mathbf{W}^+ = E[\mathbf{x}\mathbf{g}(\mathbf{W}^T\mathbf{x})] - E[\mathbf{g}'(\mathbf{W}^T\mathbf{x})]\mathbf{W} + \alpha(\mathbf{W}^* - \mathbf{W})$ 를 계산한다.

(3) $\mathbf{W}^* = \mathbf{W}^+ / \|\mathbf{W}^+\|$ 를 계산한다.

(4) $\|\mathbf{W} - \mathbf{W}^*\| > \epsilon$ 이고 $\|\mathbf{W} + \mathbf{W}^*\| > \epsilon$ 이면 식 (4)에 따라 \mathbf{W} 를 계산하고, 그렇지 않으면 단계 4를 종료한다.

단계 5 : 구해진 역혼합행렬 \mathbf{W} 로부터 원 영상, 워터마크, 키 영상을 구성한다.

단계 6 : 필요에 따라 추출된 워터마크의 인증을 수행한다.

3. 시뮬레이션 및 결과분석

제안된 워터마킹의 타당성과 성능을 평가하기 위해 레나(lena) 원 영상과 시험패턴의 키 영상을 이용하여 실험하였다. 여기서 워터마크로는 문자영상을 이용하였다. 추출된 영상, 워터마크, 그리고 키 영상의 성능을 평가하기 위해 정규상관계수(normalized correlation coefficient : ncc)와 절대평균오차(absolute mean error : ame)를 각각 정의하여 비교하였다. $ncc = \frac{\sum_{i=1}^{N^2} x(i)y(i)}{\sqrt{\sum_{i=1}^{N^2} x(i)^2 \sum_{i=1}^{N^2} y(i)^2}}$ 이며, $ame = \frac{1}{N^2} \sum_{i=1}^{N^2} |x_i - y_i|$ 이다. 여기서 x_i 와 y_i 는 각각 원 신호와 추출신호의 i 번째 값이며, N^2 은 신호의 개수이다. 시뮬레이션은 펜티엄IV-2.2G 컴퓨터 상에서 Matlab 5.1로 실행하였다.

그림 1은 256×256 픽셀의 영상들을 나타낸 것이다. 그림 1(a)와 (b)는 각각 워터마크 삽입전의 원 영상, 키 영상, 워터마크(문자)와 그들의 선형결합에 의해 생성된 워터마크가 삽입된 영상을 각각 나타낸 것이다. 그림 1(b)의 워터마크된 영상들에서는 워터마크와 키 영상에 대한 정보를 알 수 없으며, 워터마크의 삽입에 따른 원 영상의 화질저하도 보이지 않는다. 그림 1(c)는 그림 1(b)에서 평균이 0이고 분산이 0.0000427인 가우스 잡음을 첨가된 워터마크된 영상이다. 이는 원래 워터마크된 영상이 외부로부터

잡음의 공격을 받은 경우이다. 실험에서는 좀 더 효과적인 ICA를 위해 전처리과정으로 그림 1(c) 영상들의 영평균과 whitening을 수행하였다.

그림 2는 그림 1(c)의 잡음이 첨가된 워터마크된 영상들로부터 뉴우턴법의 FP_ICA에 의하여 추출된 영상들이다. 그림에서 보면 영상들의 추출순서가 두 기법 상호간 및 원래 영상과 일치하지 않음을 알 수 있다. 이는 ICA에서 독립성분을 추출할 때 순서를 고려하지 않는 속성이 반영되었기 때문이다. 또한 원래 영상들과 추출된 영상들을 순서에 맞추어 서로 비교하면 육안으로 상당히 정확하게 추출되었음을 알 수 있다. 추출된 영상들은 각각 워터마크된 영상의 독립성분들이며, 여기서 원본 영상들을 분리하는데 소요되는 알고리즘의 총 반복수는 18회 소요되었다. 한편 원본 영상들의 추출을 위해서 여기서는 별도의 원 영상이나 키 영상, 워터마크 등에 대한 정보를 요구하지 않는다.

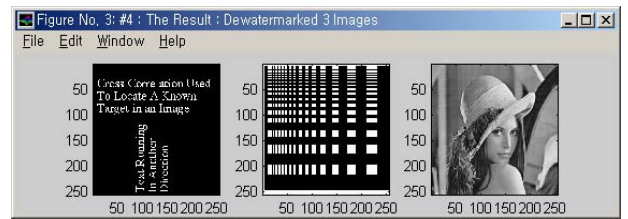
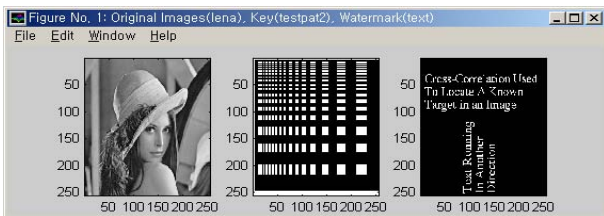


그림 2. FP-ICA에 의해 추출된 영상들



(a) 원 영상, 키 영상, 워터마크(문자영상)



(b) 워터마크된 영상



(c) 잡음이 첨가된 워터마크된 영상
(평균=0, 분산=0.0000427)

그림 1. 256×256 픽셀의 워터마킹 영상들

표 1은 워터마크된 영상에 잡음을 첨가한 영상들을 대상으로 뉴우턴법의 FP-ICA를 시행하고, 그에 따른 학습반복수 n , CPU 시간인 추출시간 t , 절대평균오차 ame , 그리고 정규상관계수 ncc 를 나타낸 것이다. 여기서 정규상관계수의 경우는 $-1 \leq ncc \leq 1$ 의 범위로 $-$ 는 원래 영상과 추출된 영상사이의 부호가 반대임을 나타내며, 1은 완전한 일치이고 0은 상관성이 전혀 없음을 나타낸다. 추출결과로부터 3개의 원래 영상들을 모두 분리하는데 소요되는 총 학습반복수는 최소 약 15회에서 최대 약 32회이며, 추출시간은 최소 약 3.5초에서 최대 약 7.01초 정도 소요됨을 알 수 있다. 이는 잡음의 공격이 심할수록 추출을 위한 소요시간이 오래 걸림을 보여 준다. 한편 원 영상의 경우는 평균이 0이고 분산이 0.00427 정도의 가우스잡음에 대해서는 각각 0.980482 정도의 높은 정규상관계수와 0.157761 정도의 낮은 절대평균오차를 가져 잡음공격에 대해 강인함을 알 수 있다. 또한 키 영상의 경우는 분산이 0.001까지는 강건하나 잡음의 양이 많은 0.00427의 분산에서는 급격하게 약한 특성을 보임을 알 수 있다. 하지만 워터마크 영상의 경우는 가우스잡음의 분산이 0.0001이상

에서는 급격하게 정규상관계수는 떨어지고 절대평균 오차는 증가하여 다른 영상에 비해 상대적으로 잡음의 공격에 매우 취약한 특성이 있음을 알 수 있다. 일반적으로 워터마크의 추출에서 정규상관계수가 0.7이상이면 우수한 추출성능이 있는 것으로 알려져 있다[2].

표 1. FP_ICA에 의한 원본영상 추출성능

분산, σ (평균=0)	원 영상		키 영상		워터마크		학습 반복수, n (원영상, 키 영상, 워터 마크)	추출 시간, t
	정규 상관 계수, ncc	절대 평균 오차, ame	정규 상관 계수, ncc	절대 평균 오차, ame	정규 상관 계수, ncc	절대 평균 오차, ame		
0	0.998 419	0.047 579	0.999 952	0.004 593	1	0	15 (1,9,5)	3.5
0.00001	0.999 741	0.015 871	0.994 355	0.084 943	0.849 271	0.450 425	15 (10,4,1)	3.37
0.0000427	0.998 21	0.051 427	0.980 882	0.155 966	0.775 699	0.585 102	18 (4,13,1)	4.06
0.0001	0.997 966	0.053 259	0.957 155	0.233 934	0.445 116	0.896 176	17 (7,9,1)	3.84
0.000427	0.996 528	0.067 155	0.847 869	0.440 318	0.025 323	1.001 738	17 (6,10,1)	3.83
0.001	0.994 227	0.085 795	0.720 711	0.597 177	0.013 298	1.008 148	17 (8,8,1)	3.78
0.00427	0.980 482	0.157 761	0.446 626	0.844 052	0.006 778	1.007 589	32 (11,20,1)	7.01

한편 표 2는 원본 영상들 상호간의 정규상관계수를 측정한 것이다. 여기서 원 영상과 키 영상 및 키와 워터마크 사이의 정규상관계수는 음의 값을 가져 이는 두 영상 상호간의 부호가 서로 반대임을 의미한다. 또한 원 영상과 워터마크 사이 및 키와 워터마크 사이는 매우 작은 정규상관계수를 가지므로 이들 사이에는 거의 상관성이 없다. 따라서 원 영상 및 키 영상과 상관성이 적은 워터마크를 이용하는 것이 보다 우수한 워터마크 추출성능을 얻을 수 있음을 확인할 수 있다.

표 2. 원본 영상들 상호간의 정규상관계수

watermark	원 영상과 key 영상	원 영상과 watermark	key와 watermark
문자	-0.052408	0.018285	-0.009755

4. 결론

본 논문에서는 고정점 알고리즘 독립성분석을 이용하여 잡음에 강건한 디지털영상의 워터마킹을 제안하였다. 여기서 고정점 알고리즘은 워터마킹의 추출 과정에서 빠른 추출과 잡음에 강건한 개선된 추출성능을 얻기 위함이고, 독립성분석의 이용은 추출과정에서 워터마크의 위치나 크기, 원본과 키 영상 등에

대한 사전 지식의 요구를 없애기 위함이다.

제안된 기법을 256x256 픽셀의 레나 원 영상, 키 영상, 그리고 문자 워터마크에 적용한 결과, FP_ICA 워터마킹은 원본 영상의 빠른 추출과 잡음에 강건한 추출성능이 있음을 확인하였다. 특히 영상들 상호간에 상관성이 적은 워터마크가 우수한 추출성능이 있음도 확인하였다. 한편 워터마크의 검출 및 추출 과정에 원본 영상들에 대한 사전지식도 요구되지 않았다.

향후 다양한 ICA의 이용과 압축이나 기하학적 변환 등의 다양한 공격에 대한 강인성 등에 대한 연구가 이루어져야 할 것이다.

참고문헌

[1] F. Hartung and M. Kutter, "Multimedia Watermarking Techniques", *Proc. of IEEE.*, Vol.87, No.7, pp.1079-1107, July 1999

[2] G. Voyatzis, N. Nikolaidis and I. Pitas, "Digital Watermarking : An Overview", *Proc. of EUSIPCO'98*, Rhodes Greece, pp.8-11, sep. 1998

[3] J. Liu, X. Zhang, J. Sun, and M. A. Lagunas, "A Digital Watermarking Scheme Based on ICA Detection", *ICA2003*, Nara Japan, April 2003

[4] 원치선, "디지털 영상의 저작권 보호", *정보과학회지*, 제 5권, 제 12호, pp.22-27, 1997, 12

[6] P. Comon, "Independent Component Analysis -A New Concept?," *Signal Processing*, vol.36, No.3, pp.287-314, Apr. 1994

[7] T. W. Lee, 'Independent Component Analysis : Theory and Applications', Kluwer Academic Pub., Boston, 1998

[8] A. Hyvarinen and E. Oja, "A Fast Fixed Point Algorithms for Independent Component Analysis", *Neural Computation*, 9(7), pp. 1483-1492, Oct. 1997

[9] Y. H. Cho and Y. S. Park, "Independent Component Analysis of Fixed-point Algorithm Based on Secant Method", *Proc. of MLMTA'03*, Las Vegas, USA., June 2003

[10] K. Atkinson, 'Elementary Numerical Analysis', John Wiley & Sons, Inc., New York, 1993