

움직임 벡터의 정규화 및 에지의 패턴 분석을 이용한 복수 영상 기반 초해상도 영상 생성 기법

권순찬*, 유지상°

Multi-Frame-Based Super Resolution Algorithm by Using Motion Vector Normalization and Edge Pattern Analysis

Soon-chan Kwon*, Jisang Yoo°

요 약

본 논문에서는 움직임 벡터의 정규화 및 윤곽선(edge)의 패턴 분석을 이용한 새로운 복수영상 기반의 초해상도(super resolution) 영상 생성 기법을 제안한다. 기존의 복수영상 기반의 초해상도 기법의 경우 입력 동영상을 구성하는 각 영상 간 부화소(sub-pixel) 단위의 움직임과 병진이동(global translation)만이 발생한다고 가정하여 기법의 적용이 제한적이다. 또한 이러한 제한에 강한 단일영상 기반의 초해상도 영상 생성 기법의 경우 보간 시 사용할 수 있는 정보량이 제한적이라는 단점이 있다. 본 논문에서는 기존 기법의 단점인 부화소 단위의 움직임에 대한 제한을 움직임 벡터의 정규화 기법을 통해 해결하고, 윤곽선 패턴 분석을 기반으로 한 2*2 블록 단위의 움직임 추정을 통해 병진이동에 대한 제한을 해결하였다. 또한 실험을 통하여 제안하는 기법이 기존의 이중선형(bi-linear)보간법, 단일영상과 복수영상 기반 초해상도 기법보다 우수하다는 것을 확인하였다.

Key Words : super resolution, interpolation, sub-pixel, motion estimation, kernel regression

ABSTRACT

In this paper, we propose multi-frame based super resolution algorithm by using motion vector normalization and edge pattern analysis. Existing algorithms have constraints of sub-pixel motion and global translation between frames. Thus, applying of algorithms is limited. And single-frame based super resolution algorithm by using discrete wavelet transform which robust to these problems is proposed but it has another problem that quantity of information for interpolation is limited. To solve these problems, we propose motion vector normalization and edge pattern analysis for 2*2 block motion estimation. The experimental results show that the proposed algorithm has better performance than other conventional algorithms.

1. 서 론

영상의 해상도를 확대하기 위해 기존 화소들 사이에 적절한 값을 삽입하는 기술을 영상 보간법(interpolation)이라 한다. 디지털 카메라 혹은 휴대 기기에 내장되어 있는 카메라와 같이 디지털 영상

획득 매체의 증가와 매체간의 성능 차이로 인하여 저해상도 영상을 보완하는 영상 보간법의 중요성은 점점 더 커지고 있다. 기존의 전통적인 영상 보간법은 저해상도 영상 화소의 가중치 합(weighted-sum)에 기반을 두고 있으며, 그 중 대표적인 방법으로 최단입점(nearest neighborhood) 보간법, 이중선형

* 본 연구는 한국연구재단 기초연구사업(2011-0006791) 지원을 받아 수행되었습니다.

• 주저자 : 광운대학교 전자공학과 디지털 미디어 연구실, honest_soonc@kw.ac.kr, 학생회원

° 교신저자 : 광운대학교 전자공학과 디지털 미디어 연구실, jsyoo@kw.ac.kr, 중신회원

논문번호 : KICS2012-07-313, 접수일자 : 2012년 7월 10일, 최종논문접수일자 : 2013년 1월 7일

(bi-linear) 보간법, 고등차수(bi-cubic) 보간법 등이 있다¹¹. 그러나 이러한 방법들은 일종의 저주파 통과 필터(low pass filter)의 개념으로, 필연적으로 영상 내 경계부분이 모호해지는 번짐(blurring) 현상이 나타난다.

이와 같은 문제를 해결하기 위해 다양한 초해상도(super resolution) 영상 생성 기법의 연구가 진행된 바 있다. 초해상도 영상 생성 기법은 동영상에서 1초에 재생되는 영상의 수가 통상 20에서 30장으로 서로 상관도가 매우 높은 특성을 이용하여, 목표 영상을 기준으로 연속된 복수의 저해상도 영상의 정보를 이용하여 하나의 고해상도 영상을 얻는 기술이다². 또한 최근에는 그 의미가 확장되어 단일영상을 사용한 초해상도 영상 생성 기법 또한 다양하게 제안되었다. 그 중 하나로 입력 영상에 이산 웨이블릿 변환 (Discrete Wavelet Transform; DWT)을 이용한 방법이 제안되었다³. 이 기법은 이산 웨이블릿 변환 시 거쳐야 하는 다운 샘플링(down-sampling) 과정을 생략하는 것을 가장 큰 특징으로 한다. 이를 통해 원본 영상과 같은 크기의 저주파-저주파(LL), 저주파-고주파(LH), 고주파-저주파(HL), 고주파-고주파(HH) 부대역(sub-bands)들을 얻은 후 LL 부대역을 제외한 나머지 고주파 부대역에 다양한 처리를 한 뒤 역 변환을 하면, 고주파 성분이 잘 보존된 확대 영상을 얻을 수 있다. 그러나 이 방법은 전통적인 보간법과 마찬가지로 외부의 추가적인 정보를 사용하지 않는다는 점에서 성능이 제한적이다.

또한 기존에 제안된 복수영상 기반의 초해상도 영상 생성 기법들의 경우, 초해상도 영상 생성 기법이 적용되는 입력 영상이 단순 회전(rotation)이나 전역이동과 같은 병진이동(global translation)과, 매우 작은 부화소(sub-pixel) 단위의 움직임만을 포함하고 있어야 좋은 결과를 얻을 수 있고^{4,5}, 이러한 제한을 극복하기 위해 몇 가지 다른 기법이 제안되었으나^{6,7} 이 기법들의 경우 객체의 내부 질감(texture)을 모호하게 하는 문제점이 있다. 다른 방법으로는 한 장의 고해상도 영상을 핵심영상(key-frame)으로 하여 이 정보를 기반으로 인접한 영상들에 적용하는 초해상도 영상 생성 기법이 제안되었으나⁸, 이 방법은 동영상 압축에 쓰이는 기법으로, 핵심영상으로 사용할 최소 한 장의 고해상도 영상이 존재하지 않으면 적용할 수 없다는 단점이 있다.

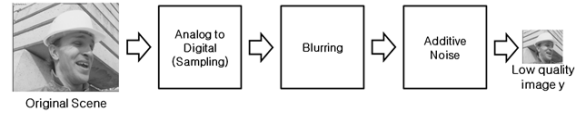


그림 1. 저해상도 영상이 생성되는 과정
Fig. 1. Generation of low resolution images

본 논문에서는 이러한 문제점들을 극복하기 위하여 움직임 벡터의 정규화 및 윤곽선(edge)의 패턴 분석을 이용한 새로운 복수 영상 기반의 초해상도 영상 생성 기법을 제안한다. 공통적으로 초해상도 영상 생성 기법은 그림 1에 보인 것처럼 한 장의 저해상도 영상의 생성과정을 가정한다. 초해상도 영상 생성 기법은 그림에 나타난 것처럼 어떤 자연 영상(nature scene)이 디지털 신호화 되는 과정에서 다운 샘플링 되고, 센서의 오작동 등으로 인하여 잡음이 더해져 하나의 저화질 영상이 생성된다고 가정하고 이 문제를 역으로 해결하는 과정으로 설명된다².

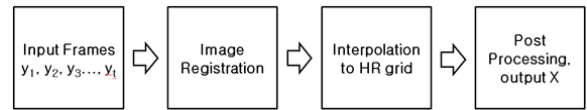


그림 2. 복수 영상 기반 초해상도 영상 생성 기법의 흐름도
Fig. 2. Block diagram of multi-frame based super resolution

일반적으로, 복수 영상 기반의 초해상도 영상 생성 기법들이 공통적으로 가지는 흐름도는 그림 2에 보인 것과 같다. 첫 번째 단계로 초해상도 영상 생성 기법을 적용하려는 목표 영상을 포함하여 총 t 장의 영상이 입력된다. 이 경우, 저해상도 영상 y_1 부터 y_t 는 인접한 영상이어야 하며, 장면의 전환이 발생해서는 안 된다. 이렇게 정해진 입력영상들은 목표 영상을 기준으로 그 위치관계를 정합해야 한다. 그 후 위치관계를 기반으로 하여 얻어진 정합점들은 고해상도 영상의 화소 값을 추정하는 데 사용한다. 이러한 결과에 최종적으로 블록현상 제거 필터(de-blocking filter) 혹은 초점복원 필터(de-blurring filter) 적용 등을 후처리 과정으로 하여 한 장의 고해상도 영상 X 를 얻게 된다⁷. 본 논문에서는 그림 2의 과정 중 2번째와 3번째 단계인 정합과 고해상도 보간의 내용을 주로 다루며, 후 처리에 관한 내용은 다루지 않는다.

본 논문의 구성은 다음과 같다. II장에서는 제안

하는 기법을 구성하는 각 단계에 관한 자세한 설명을 하고 III장에서는 다양한 실험을 통하여 제안하는 기법의 우수성을 보인다. 마지막으로 IV장을 통한 결론으로 끝을 맺는다.

II. 제안하는 기법

본 장에서는 제안하는 복수 영상 기반의 초해상도(super resolution) 영상 생성 기법을 이루는 각 단계들을 세부적으로 설명한다. 우선 2.1절에서는 부화소(sub-pixel) 단위의 화소 이동에 관한 개념을 설명하고, 2.2절에서 6-tap FIR(finite impulse response) 필터를 통한 부화소 단위의 움직임 추정(motion estimation)에 관한 설명을, 2.3절에서는 움직임 추정으로 얻어진 움직임 벡터의 정규화 기법에 관한 설명을 한다. 2.4절에서는 보다 효과적인 정합점의 참조를 위해 제안하는 블록단위 윤곽선(edge)의 패턴분석 기법을 설명하고, 2.5절을 통해 본 논문에서 사용되는 1차원 커널 추정법(kernel estimation)에 대하여 간략하게 설명한다. 마지막으로 2.6절에서 제안하는 기법의 적용 과정을 정리하여 설명하는 것으로 II장을 마무리 한다.

2.1. 부화소 단위의 화소이동

인간이 육안으로 보는 자연 영상은 연속적인 아날로그 신호이다. 영상신호를 컴퓨터 처리가 가능한 디지털 신호로 변환하기 위해서는 음성신호와 같이 신호를 디지털화 하는 과정이 필요하다. 그런데 임의의 자연 영상(nature scene)이 그림 3에 보인 것처럼 저해상도 영상이 되는 과정에서 영상획득 기기의 센서 간 샘플링 간격으로 인한 필연적인 앨리어싱(aliasing)이 발생한다. 이로 인해 원 영상신호의 고주파 성분이 손실된다⁹⁾.

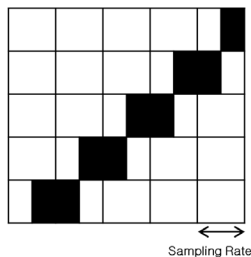


그림 3. 디지털 변환(ADC) 과정에서 발생하는 앨리어싱
Fig. 3. Aliasing caused by analog to digital process

초당 20에서 30장의 영상으로 구성된 동영상에서

현재의 영상을 $f(t)$ 라 하고, 바로 전의 영상을 $f(t-1)$ 라고 가정하면 장면의 전환 부분이 아닌 경우 두 영상은 서로 근소한 차이만을 가질 것이다. 그림 4는 영상 간 존재하는 객체의 움직임을 보인다. 그림 4에서 왼쪽 열은 $f(t-1)$ 영상, 오른쪽 열은 $f(t)$ 영상을 의미한다⁴⁾.

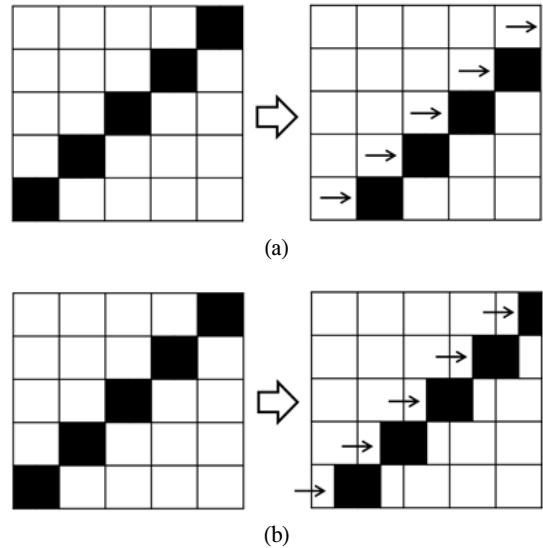


그림 4. 화소이동 (a) 정수단위 (b) 부화소단위, $f(t-1)$ (첫 번째 열), $f(t)$ (두 번째 열)
Fig. 4. Pixel movement (a) integer unit (b) sub-pixel unit, $f(t-1)$ (1st col.), $f(t)$ (2st col.)

그림 4의 (a)는 영상 간 객체의 이동이 샘플링 간격인 정수단위로 이동한 경우를 보인다. 이 경우 디지털화 된 영상은 과 의 해당 객체가 같은 윤곽선 정보를 가지게 된다. 그림 4의 (b)는 영상 간 객체가 샘플링 간격보다 작은 부화소 단위로 이동한 경우이다. 이 경우 첫 번째 경우와 달리 두 영상은 같은 객체에 대하여 서로 다른 윤곽선 정보를 가지게 된다²⁾.

첫 번째 경우처럼 복수의 영상이 서로 같은 윤곽선 정보를 가지고 있다면, 해당 영상들로는 초해상도 영상 생성 기법 적용이 불가능하고 이 경우 영상 간에 정수 단위의 화소이동이 일어났다고 정의한다. 그러나 두 번째 경우는 서로 다른 윤곽선 정보를 조합하여 초해상도 영상 생성 기법을 적용할 수 있으며 이 경우 영상 간에 부화소 단위의 화소 이동이 일어났다고 정의한다⁴⁾. 따라서 복수 영상을 기반으로 하는 초해상도 영상 생성 기법을 적용하기 위해서는 입력 영상 간 부화소 단위의 이동이 보장되어야 한다.

2.2. 6-tap FIR 필터를 통한 부화소 단위의 움직임 추정

2.1절에서 언급한 바와 같이 복수 영상 간 초해상도 영상 생성 기법이 가능하기 위해서는 입력영상들이 부화소 단위의 움직임을 가져야 하며, 부화소 단위의 움직임을 찾아내기 위해 입력영상들에 대한 부화소 확장이 필요하다. 본 논문에서는 1/4 화소(quarter-pixel) 단위의 움직임 추정을 통한 정합방법을 사용하기 때문에, 입력 영상들에 대하여 가로 4배, 세로 4배로 총 16배의 영상 확장이 필요하다. 여기에 사용되는 방법은 대표적으로 이중선형(bi-linear), 고등차수(bi-cubic), 란초스(Lanczos) 보간법 등이 있다⁴⁴. 이러한 부화소 확장 방법들 중 움직임 추정 및 초해상도 영상 생성 기법에 가장 적합한 보간법은 H.264/AVC의 표준에서 사용되는 6-tap FIR 필터로서 기존 기법에서 실험적으로 증명되었다⁴⁴. 따라서 본 논문에서는 6-tap FIR 필터 기반의 부화소 확장을 적용한다. 식 (1)은 1/2 화소(half-pixel) 탐색 과정에 사용된다¹¹⁰.

$$\begin{aligned}
 b_1 &= (E - 5F + 20G + 20H - 5I + J) \\
 h_1 &= (A - 5C + 20G + 20M - 5R + T) \\
 b &= (b_1 + 16) \gg 5 \\
 h &= (h_1 + 16) \gg 5 \\
 j_1 &= cc - 5dd + 20h_1 + 20m_1 - 5ee + ff \\
 j &= (j_1 + 512) \gg 10
 \end{aligned} \tag{1}$$

여기서 b 와 h 는 각각 1/2 화소 단위의 화소 값을 의미하고, $A, C, G, M, R, T, E, F, H, I, J$ 는 각각 정수단위의 원본 화소 값을 의미한다.

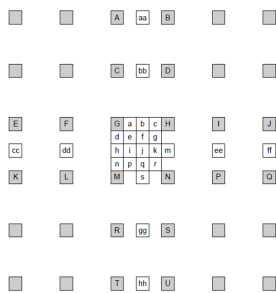


그림 5. 6-tap FIR 필터 설명을 위한 화소 배치도
Fig. 5. Pixel chart for the description of 6-tap FIR filter

또한, 각 정수단위의 화소에 곱해지는 숫자는 가중치를 나타낸다. 그림 5에 식 (1)에 따른 화소의 배치도를 보인다. 그림 5에서 대문자로 표시되어 있

는 회색 사각형은 원본 정수 화소를 의미한다. 그림 5과 같이 1/2화소를 구한 후 1/4 화소(quarter-pixel) 또한 식 (2)와 같이 선형 보간법을 적용해 구할 수 있다¹¹⁰.

$$a = (G + b + 1) \gg 1 \tag{2}$$

식 (1)과 식 (2)를 이용하여 가로 4배와 세로 4배, 총 16배로 부화소로 확장된 입력영상들을 가지고 목표 영상에 대하여 블록단위의 움직임 추정을 적용하게 된다. 움직임 추정은 원래 영상압축을 통해 제안된 기법으로 그림 6과 같이 현재 영상과 다음 영상의 시간적 중복성(temporal redundancy)을 이용하여 이전영상의 차분(difference)만을 저장하여 현재영상을 만들어 내는 압축 원리를 가진다¹¹¹.

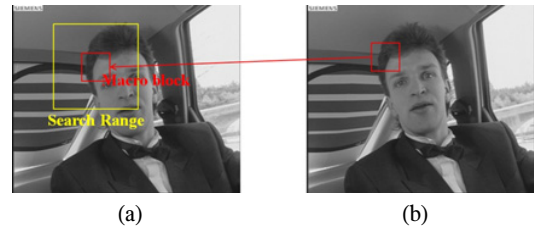


그림 6. 움직임 추정에서의 영상 간 시간적 중복성 이용
(a) $(t-1)^{\text{th}}$ 영상 (b) t^{th} 영상

Fig. 6. Using property of temporal redundancy of motion estimation (a) $(t-1)^{\text{th}}$ frame (b) t^{th} frame

본 논문에서는 2×2 블록단위의 움직임 추정을 이용함으로써 영상 간 복잡한 움직임이나 지역적인 움직임(local motion)에 강한 영상 정합이 가능하다. 영상 간 유사블록의 탐색 기준은 식 (3)의 SAD(sum of absolute difference)값을 이용한다¹¹¹.

$$SAD(i, j) = \sum_{y=0}^1 \sum_{x=0}^1 |B_i(x, y) - B_p(i+x, j+y)| \tag{3}$$

여기서 $SAD(i, j)$ 는 탐색영역 내 (i, j) 위치의 SAD값을 나타내고, x 와 y 는 처리 블록 내의 좌표를 나타내며, 0과 1의 값을 가질 수 있다. B_i 는 목표 영상내의 처리 블록, B_p 는 p 번 입력 저해상도 영상내의 처리 블록을 나타낸다. 움직임 추정은 많은 연산을 필요로 하는 과정으로, 그 속도를 높이기 위해 다양한 방법이 제안되었으나^{12,13} 모두 탐색의 정확도를 어느 정도 감소시킨다. 움직임 추정의 고속화는 본 논문의 범위가 아니며, 초해상도 영상 생

성 기법의 결과는 정합의 정확도에 의해 크게 좌우되므로 본 논문에서는 전역탐색(full search) 방식의 움직임 추정을 사용한다.

2.3. 움직임 벡터의 정규화

2.2절의 과정을 통해 영상 간의 움직임 벡터를 찾은 후 이를 정규화 하는 과정이 필요하다. 2.1절에서 언급한 바와 같이 복수 영상을 기반으로 하는 초해상도 영상 생성 기법은 입력영상 간 목표 영상에 대하여 부화소 단위의 이동이 보장되어야 한다^[2]. 즉, 영상 내 객체의 움직임은 반드시 한 화소 이내의 부화소 범위에서 나타나야 한다. 그러나 이러한 제한사항은 초해상도 영상 생성 기법을 적용할 수 있는 영상을 상당히 제한적으로 만든다. 다시 말해 움직이지 않는 사물이나 배경을 하나의 영상 매체로 얻을 경우 쉽게 이러한 영상을 획득할 수 있으나, 객체의 움직임이 복잡하거나 클 경우 기존의 초해상도 영상 생성 기법을 적용하기 힘들다. 실제로 일반적인 동영상은 이러한 조건을 충분히 만족하지 않는다. 따라서 본 논문에서는 움직임 벡터의 정규화를 통해 한 화소 이상의 큰 움직임 벡터를 갖는 일반적인 경우에 대해서도 초해상도 영상 생성 기법을 적용할 수 있도록 한다. 그림 7은 이에 따른 움직임 벡터 정규화 기법의 흐름도이다.

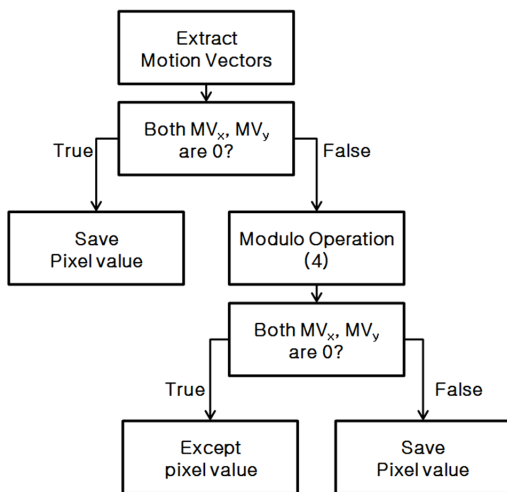


그림 7. 움직임 벡터 정규화 기법의 흐름도
Fig. 7. Block diagram of motion vector normalization algorithm

먼저 전체 p 장의 입력영상에서 목표 영상 내 하나의 처리 블록에 대하여 나머지 $(p-1)$ 장에 대한 블록 움직임 추정을 한다. 움직임추정 결과로 한 개의 목표 보간점에 대하여 총 $2*2*(p-1)$ 개의 움직

임 벡터를 얻을 수 있다. 이 벡터가 0의 값을 가지는 경우 목표 보간점과의 거리가 0이므로 정규화를 거치지 않고 그 값을 그대로 사용한다. 만약 0의 벡터 값을 가지는 화소가 복수 개라면, 해당 화소들의 값을 평균하여 목표 보간점에 대입한다. 반대로 추출한 움직임 벡터의 거리가 0이 아닌 경우, 4의 나머지 연산(modulo)으로 움직임 벡터를 정규화한다. 여기서 4의 나머지 연산을 하는 이유는 1/4화소 정확도의 움직임 추정을 위해 입력영상들에 대하여 가로 4배, 세로 4배의 부화소 확장을 적용하였기 때문이다. 만약 나머지 연산의 결과가 0이라면, 해당 화소는 부화소 단위가 아닌 정수단위의 이동량을 가지므로 목표점 보간에 사용하지 않는다. 나머지 연산의 결과가 0인 경우를 제외하고 4의 나머지 연산의 결과로 가능한 값은 1, 2, 3으로, 각각에 대하여 0.25, 0.5, 0.75의 부화소 이동량을 부여하고 해당 화소 값을 저장한다. 이렇게 얻어지는 정규화된 움직임 벡터가 가리키는 각 입력 저해상도 영상의 정합점들과 그 화소 값들은 2.5절에서 언급할 커널 추정법(kernel estimation)을 통한 목표 화소 보간에 사용된다.

2.4. 윤곽선의 패턴 분석

기존 복수 영상 기반의 초해상도 영상 생성 기법들의 경우 정합의 결과로서 얻은 화소 값들에 대해 정합의 정확도나 윤곽선의 패턴을 고려하지 않고 다양한 보간법을 통해 목표 보간점의 화소 값을 구한다. 그러나 이 부분을 고려하지 않으면 적합하지 않은 화소 값이라 해도 부화소 이내의 움직임 조건만 만족하면 사용될 수 있어 제안하는 기법의 결과 품질을 저하시킨다. 이 오류를 그림 8에 보인다. 그림 8의 (a)에 보인 것처럼 검게 칠해진 부분은 윤곽선을 나타내며, 이 윤곽선이 포함된 블록을 초해상도 복원 시 사용하기 위한 정합점들 중 일부를 붉은색과 파란색으로 나타내었다. 파란색으로 표시된 위치에 속해있는 정합점들은 사용되기에 적합하지만, 붉은색으로 표시된 위치의 정합점들을 사용하면 화소 값의 정확도가 떨어지고, 경우에 따라 화소 값이 튀는 격자현상이 나타날 수 있다. 이러한 오류를 개선하기 전과 개선한 후의 결과를 그림 8의 (b)와 (c)에 각각 보이고 붉은 원으로 표시하였다.

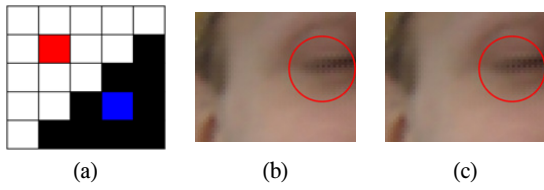


그림 8. 윤곽선의 패턴을 고려하지 않은 정합점의 사용 (a) 다양한 정합점의 위치 (b) 패턴 분석 적용 전 (c) 패턴 분석 적용 후

Fig. 8. Use of registration points without considering pattern of edge (a) various position of registration point (b) before using pattern analysis (b) after using pattern analysis

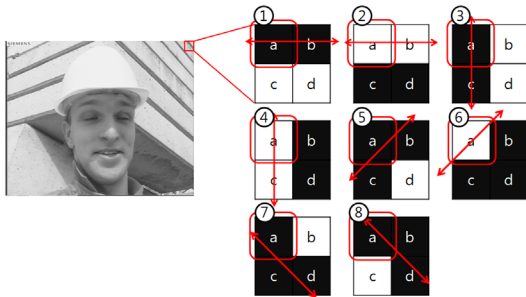


그림 9. 윤곽선의 패턴 고려를 위한 2*2 블록의 8가지 패턴

Fig. 9. Eight edge patterns of 2*2 block for the pattern analysis

본 논문에서는 2*2 블록단위의 처리를 함으로써 그림 9에 보인 것처럼 총 여덟 가지의 윤곽선 패턴 분석을 할 수 있다. 판단의 기준은 다음과 같다. 예를 들어 그림 9의 ①의 경우 a와 b, c와 d의 화소 값의 차이가 매우 근소해야 하고, a와 c, b와 d의 화소 값의 차이는 충분히 커야 한다. 이러한 차이 값 역시 문턱 값을 정해서 판별하며, 문턱 값에 따라 패턴 분석의 민감도를 조절할 수 있다. 이러한 기준에 의해 ①의 패턴으로 판단이 될 경우, 모인 정합점들에 대하여 c와 d범위 내에 있는 점들은 사용하지 않는다. 나머지 7개 윤곽선 패턴의 경우도 마찬가지로의 방식을 적용한다. 이러한 윤곽선의 패턴 분석방법 외에, 식 (4)를 이용하여 부화소 탐색으로 구해진 목표 보간점과 일정한 문턱 값 이상 차이가 나는 정합점들은 사용하지 않는다.

$$|V(P_i) - V(P_r)| < T \quad (4)$$

여기서 $V(P_i)$ 는 정합점 탐색의 기준이 되는 원본 화소 값을, $V(P_r)$ 은 각 정합점의 화소 값을 의미한다. 최적의 T 값을 통하여 잘못된 정합점의 사

용으로 나타날 수 있는 격자 현상을 해결할 수 있다. 이 절에서 설명한 두 가지 방법을 이용하여 결과영상에서 경계 부근의 잡음을 효과적으로 제거할 수 있다.

2.5. 1차원 커널 추정법

2.4절의 과정을 통해 선택된 정합점들을 이용하여 목표로 하는 보간점의 화소 값을 추정한다. 본 논문에서 사용되는 1차원의 커널 추정법(kernel estimation)은 데이터가 주어졌을 때 복수 표본의 위치와 해당 값을 근거로 하여 임의의 위치의 값을 구하는 추정법이다. 커널 추정법은 추정 식의 차수에 따라 그 정확도가 결정되나, 차수 증가에 따른 식의 복잡도가 매우 커지기 때문에 일반적으로 차수를 0으로 하여 정규화 한 추정 식을 사용한다. 식 (5)에 보인 것처럼 추정 식의 차수를 0으로 하여 정규화 된 경우 이를 nadaraya-watson 커널 추정식이라 한다^[4].

$$\widehat{f}_{NW}(x) = \frac{\sum_{i=1}^N Y_i K\left(\frac{x_i - x}{h}\right)}{\sum_{i=1}^N K\left(\frac{x_i - x}{h}\right)} \quad (5)$$

여기서 $K(u)$ 는 가우시안 기반의 커널 식을 의미하며, 식 (6)으로 나타낼 수 있다.

$$K(u) = 0.6171e^{-u^2/2}I(-1/2 \leq u \leq 1/2) \quad (6)$$

또한, 식 (6)의 $x_i - x$ 는 목표 보간점과 i 번째 정합점 사이의 유클리드 거리(euclidean distance)를 나타내고, Y_i 는 i 번째 정합점의 화소 값을, h 는 커널 추정 식의 대역폭(bandwidth)을 나타낸다. 2.3절에서 설명한 정규화 과정을 통해 얻은 움직임 벡터는 가로와 세로 방향으로 각각 0.25, 0.5, 0.75의 값을 가질 수 있다. 그림 10에 정합점들과 보간하고자 하는 목표점과의 위치관계를 보인다. 그림 10의 (a)에 보인 것처럼 각 정합점들의 목표 보간점과의 유클리드 거리와 해당 화소 값이 앞서 설명한 식 (5)에서 사용된다. 이를 통해 최종적인 목표 보간점의 화소 값을 구할 수 있다.

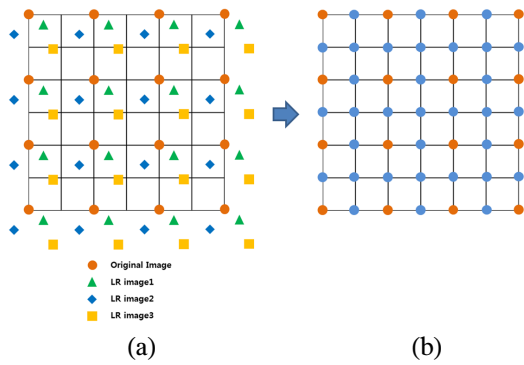


그림 10. 정합점들과 목표 보간점의 위치 (a) 보간 전 (b) 보간 후
 Fig. 10. Positions of registration and target pixels and interpolation (a) before (b) after

2.6. 전체 기법의 정리

2.1에서 2.5까지는 제안하는 기법에서 중요한 과정을 개별적으로 설명하였다. 그림 11은 각 부분을 종합하여 나타내는 전체 기법의 순차적인 흐름도이다.

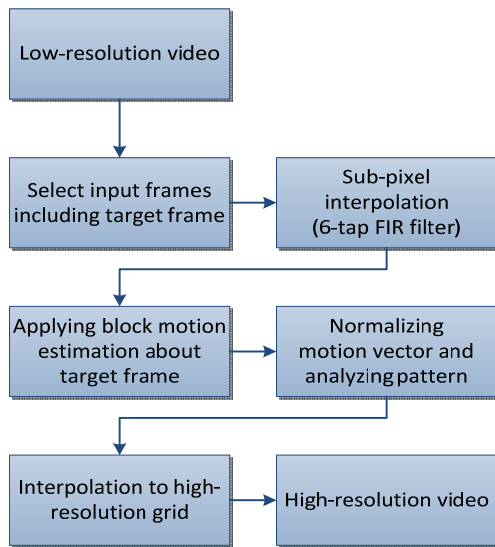


그림 11. 제안하는 기법의 흐름도
 Fig. 11. The block diagram of the proposed algorithm

우선 저해상도의 동영상이 입력으로 주어지면, 동영상을 구성하는 각 영상 단위로 제안하는 기법을 적용하게 된다. 기법을 적용하고자 하는 영상을 포함하여 전, 후 각 3장 총 7장의 영상이 사용되게 된다. 여기에 부화소 단위의 정합을 위해 6-tap FIR 필터를 이용하여 부화소 보간이 적용된다. 이 과정을 거치게 되면 각 입력 저해상도 영상들은 가로, 세로 각 4배씩 총 16배로 영상이 확장된다. 다음

로 여기에 2*2 블록 단위의 움직임 추정을 적용하여 움직임 벡터를 추출하고, 추출된 움직임 벡터는 정규화를 거치게 된다. 또한 해당 블록의 패턴이 제안한 기법에 의해 정의되면 그 결과에 따라서 정합 점들이 선택적으로 화소 값 계산에 사용된다. 마지막으로 1차원의 커널 추정식을 통하여 목표 화소의 위치와 정합점들의 위치간 거리에 따른 가중 합 (weighted-sum)으로 최종 화소 값을 계산한다. 이렇게 기법이 적용된 영상들이 모여 하나의 고해상도 동영상이 출력된다.

III. 실험 조건 및 결과

본 논문에서 제안한 기법의 성능을 확인하기 위한 실험 조건은 다음과 같다. 우선 352*288(CIF) 크기의 ‘Mother and Daughter’, ‘News’, ‘Carphone’, ‘Coastguard’, ‘Tempete’, 5개의 동영상들에 대하여 각각 연속된 30장의 영상을 176*144 (QCIF) 크기로 다운 샘플링(down-sampling) 한다. 그 후 다운 샘플링 된 영상에 각각 이중선형 (bi-linear)보간법, 이산 웨이블릿 변환 (Discrete Wavelet Transform; DWT) 기반의 초해상도 영상 생성 기법^[3], 기존에 제안된 복수 영상 기반의 초해상도 영상 생성 기법^[4]과 제안된 기법을 적용하여 다시 원본 크기로 복원 후 주관적, 객관적으로 결과를 비교하는 방식으로 실험을 진행한다. 객관적인 성능 비교를 위하여 본 논문에서는 식 (7)의 PSNR을 기준으로 사용한다^[14].

$$PSNR(f, g) = 10 \log_{10} (255^2 / MSE(f, g)) \tag{7}$$

$$where, MSE(f, g) = 1 / MN \sum_{i=1}^M \sum_{j=1}^N (f_{ij} - g_{ij})^2$$

여기서 f 는 원본 영상, g 는 비교하고자 하는 영상, M 과 N 은 각각 영상의 가로와 세로의 길이를 나타낸다. 이 수치가 높을수록 비교하고자 하는 영상 g 는 원본영상 f 에 가깝고, 기법의 성능이 우수한 것으로 정의한다.

표 1에 총 5개의 실험 영상에 따른 각각의 PSNR 측정결과를 나타내었다. 최초 30장의 실험영상을 대상으로 PSNR을 측정하고, 그 평균수치를 나타내었다. 표에서 확인할 수 있듯이 제안하는 기법을 기존 기법들과 비교했을 때, 각 실험영상에 따라 다르지만 가장 높은 PSNR 수치를 나타내는 것

을 확인할 수 있다. 특히 이중선형(Bi-linear) 보간 법과의 차이는 평균 2.5db이상으로 매우 큰 것을 알 수 있으며, ‘Mother and daughter’, ‘News’ 영상의 경우 3.8db 이상의 높은 수치차이를 나타낸다.

표 1. 각 실험 영상의 PSNR 측정결과(db)
Table 1. PSNR of each test sequences (db)

	Bi-linear	SR[3]	SR[4]	Proposed
Mother and daughter	31.32	34.83	34.32	35.03
News	24.92	27.93	27.95	28.35
Carphone	27.15	28.88	28.82	29.00
Coastguard	24.78	26.74	26.92	27.19
Tempete	22.99	24.28	24.30	24.42

그림 12부터 그림 16을 통해 실험영상에 따른 각 보간 방법의 주관적 화질 비교를 보인다. 자세한 비교를 위해 영상의 일부를 8배 확대하여 함께 보인다.

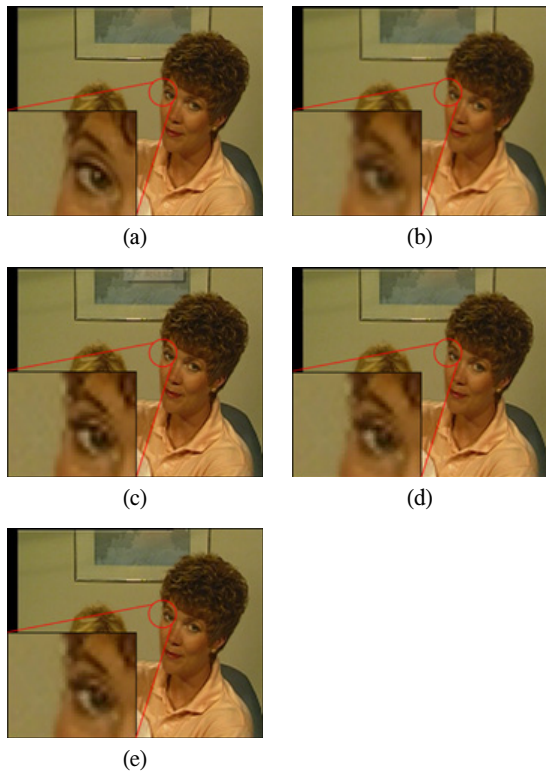


그림 12. Mother and daughter (6번째 영상) (a) 원본영상 (b) 이중선형 (c) SR[3] (d) SR[4] (e) 제안된 기법
Fig. 12. Mother and daughter(6thframe) (a) ground-truth (b) bi-linear (c) SR[3] (d) SR[4] and (e) proposed

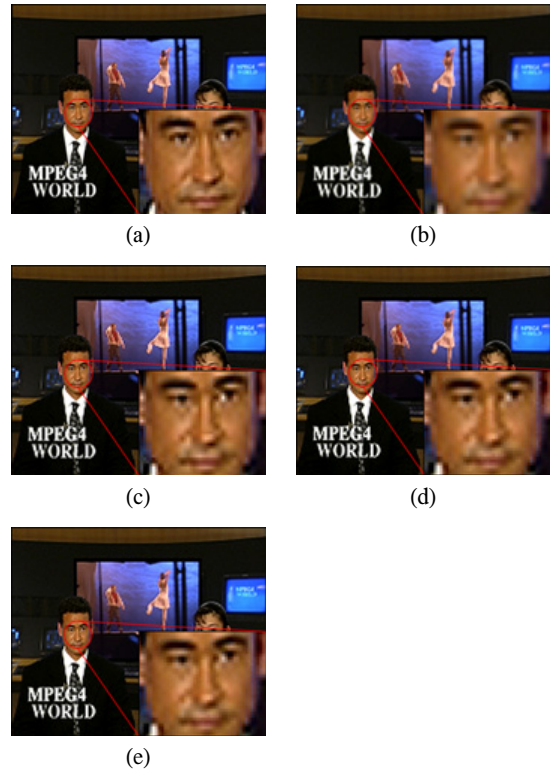


그림 13. News (8번째 영상) (a) 원본영상 (b) 이중선형 (c) SR[3] (d) SR[4] (e) 제안된 기법
Fig. 13. News(8thframe) (a) ground-truth (b) bi-linear (c) SR[3] (d) SR[4] and (e) proposed

그림 12부터 그림 16를 보면 표 1에서의 PSNR 차이와 같이 이중선형 보간법과의 선명도 차이가 매우 크다는 것을 알 수 있다. 특히 그림 12부터 그림 14의 각 확대 영상에서는 영상 내 인물의 얼굴과 경계 부근에서 기존에 제안된 방법들에 비하여 뚜렷한 윤곽선을 가진다는 것을 확인할 수 있다.

또한 그림 15의 (d)에 확대 된 객체 하단부에 2장에서 언급한 잘못된 정합점들의 사용 시 나타날 수 있는 잡음과 그것이 제안된 기법을 통해 제거된 결과를 눈으로 확인할 수 있다. 마지막으로 그림 16의 (d)에서 역시 정합오류로 인한 잡음이 관측되지만 그림 (e)에서 깨끗이 제거된 모습을 볼 수 있다. 그림 12부터 그림 16의 (c) 가 나타내는 이산 웨이블릿 변환 기반의 초해상도 영상 생성 기법^[3]의 경우 기존의 이중선형 보간법 보다는 객관적, 주관적으로 좋은 성능을 보이지만, 서론에서 언급한 것과 같이 외부의 정보를 활용할 수 없다는 점에서 화질 상승의 한계가 있다는 사실을 객관적 수치 비교를 통해 확인할 수 있다.

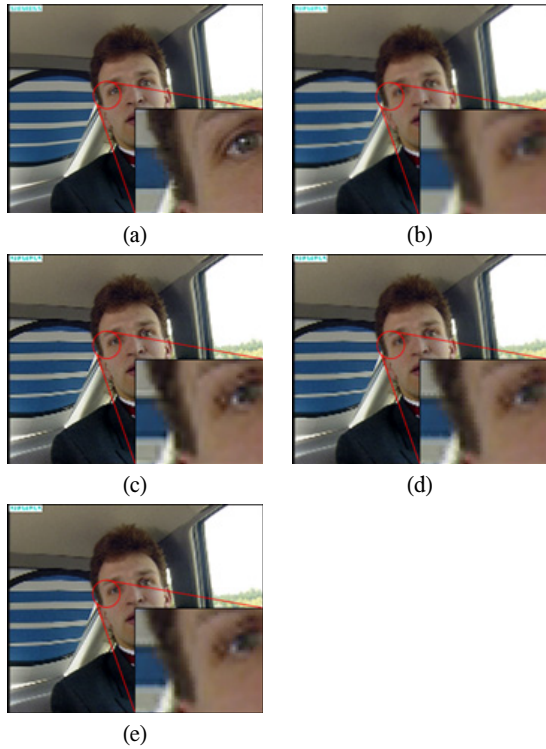


그림 14. Carphone (19번째 영상) (a) 원본영상 (b) 이중 선형 (c) SR[3] (d) SR[4] (e) 제안된 기법
 Fig. 14. Carphone(19thframe) (a) ground-truth (b) bi-linear (c) SR[3] (d) SR[4] and (e) proposed

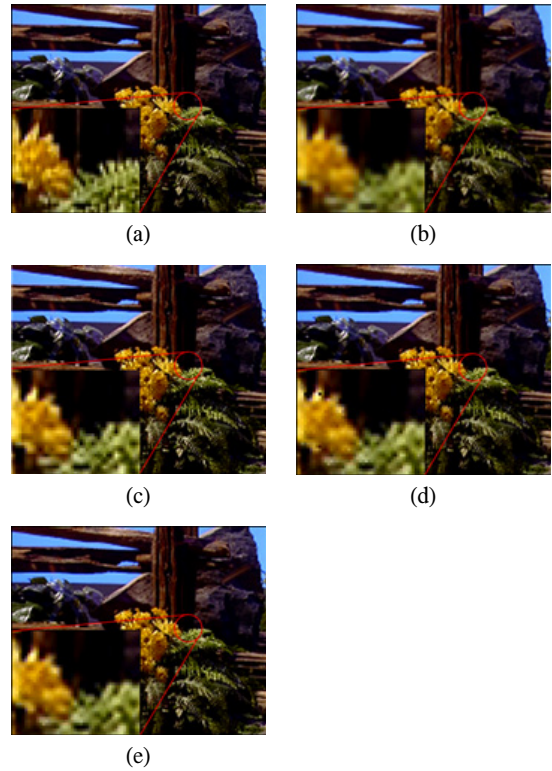


그림 16. Tempete (3번째 영상) (a) 원본영상 (b) 이중 선형 (c) SR[3] (d) SR[4] (e) 제안된 기법
 Fig. 16. Tempete(3rdframe) (a) ground-truth (b) bi-linear (c) SR[3] (d) SR[4] and (e) proposed

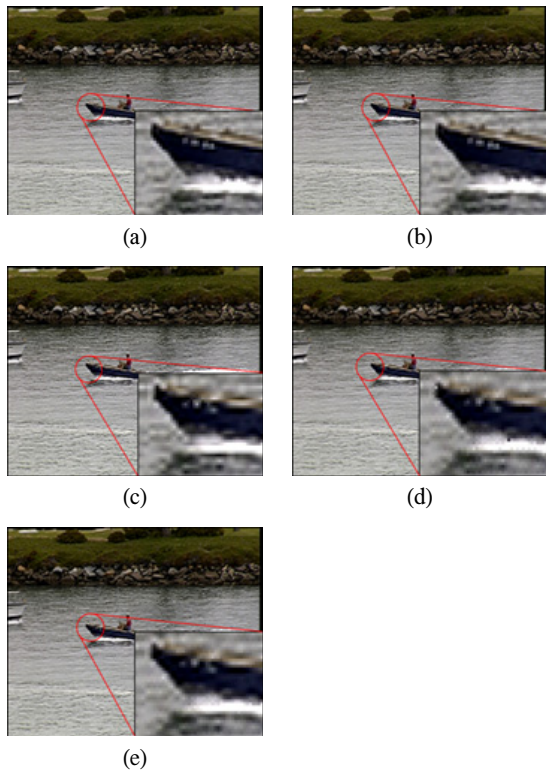


그림 15. Coastguard (4번째 영상) (a) 원본영상 (b) 이중 선형 (c) SR[3] (d) SR[4] (e) 제안된 기법
 Fig. 15. Coastguard(4thframe) (a) ground-truth (b) bi-linear (c) SR[3] (d) SR[4] and (e) proposed

IV. 결 론

본 논문에서는 기존의 단일영상 초해상도 영상 생성 기법과 복수 영상 기반의 초해상도 영상 생성 기법이 포함하고 있는 여러 제한 사항을 완화할 수 있는 새로운 초해상도 영상 생성 기법을 제안하였다. 제안하는 기법에서는 동영상 내 영상들 간에 부화소 단위의 움직임만이 존재해야 하는 제한을 움직임 벡터의 정규화를 통해 다양한 동영상에 초해상도 영상 생성 기법을 적용 가능하도록 범위를 확대하고, 윤곽선(edge)의 패턴 분석을 통하여 각 영상의 내용에 적응적인 보간을 하도록 하였다. 실험 결과 제안하는 초해상도 영상 생성 기법이 기존의 이중선형(bi-linear) 보간법, 단일영상 기반의 초해상도 영상 생성 기법과 복수 영상 기반의 초해상도 영상 생성 기법보다 객관적, 주관적으로 우수한 성능을 가진다는 것을 확인할 수 있었다.

References

- [1] G. B. Kang, Y. S. Yang and J. H. Kim, "A study on interpolation for enlarged still image," in *Proc. KIICE General Conf.*, pp. 643-648, May 2001.
- [2] S. C. Park, M. K. Park, and M. G. Kang, "Super resolution image reconstruction: a technical overview," *IEEE Signal Proc. Mag.*, vol. 20, no. 3, pp. 21-36, May, 2003.
- [3] J. M. Lim and J. Yoo, "Super resolution algorithm using discrete wavelet transform for single-image," *J. KOSBE*, vol. 17, no. 2, pp. 344-353, Mar. 2012.
- [4] S. C. Kwon and J. Yoo, "Super resolution algorithm by motion estimation with sub-pixel accuracy using 6-tap FIR filter," *J. KICS.*, vol. 37, no. 6, pp. 464-472, Jun. 2012.
- [5] V. K. Asari, M. N. Islam and M. A. Karim, "Super resolution enhancement technique for low resolution video," *IEEE trans. Consumer Elect.*, vol. 56, no. 2, pp. 919-924, May 2010.
- [6] S. Farsiu and P. Milanfar, "Kernel regression for image processing and reconstruction," *IEEE Trans. Image Process.*, vol. 16, no. 2, pp. 349-366, Feb., 2007.
- [7] M. Elad, H. Takeda and P. Milanfar, "Generalizing the nonlocal-means to super resolution reconstruction," *IEEE Trans. Image Process.*, vol. 18, no. 1, pp. 36-51, Jan. 2009.
- [8] S. C. Jeong and Y. L. Choi, "Video super resolution algorithm using bi-directional overlapped block motion compensation and on-the-fly dictionary training," *IEEE Trans. Circ. Syst. Vid.*, vol. 21, no. 3, pp. 274-285, Mar. 2011.
- [9] T. H. Kim, Y. S. Moon and C. S. Han, "Estimation of real boundary with subpixel accuracy in digital imagery," *KSPE Journals*, vol. 16, no. 8, pp. 16-22. Aug. 1999.
- [10] T. Wiegand, G. J. Sullivan, G. Bjontegaard, and A. Luthra, "Overview of the H.264/AVC video coding standard," *IEEE Trans Circ. Syst. Vid.*, vol. 13, no. 7, pp. 560-576, Jul. 2003.
- [11] N. Hirai, T. Kato, T. Song and T. Shimamoto, "An efficient architecture for spiral-type motion estimation for H.264/AVC," *IEEK General Conf.*, pp. 314-317, Jul. 2009.
- [12] Y. Ismail, J.B. McNeely, M. Shaaban, H. Mahmoud and M.A. Bayoumi, "Fast motion estimation system using dynamic models for H.264/AVC video coding," *IEEE Trans. Circ. Syst. Vid.*, vol. 22, no. 1, pp. 28-42, Jan. 2012.
- [13] H. M. Wong, O. C. Au, A. Chang, S.K. Yip and C.W. Ho, "Fast mode decision and motion estimation for H.264(FMDME)," *IEEE Int. Symp. Circ. and Syst.*, pp. 21-24, May 2006.
- [14] S. Winkler, "The evolution of video quality measurement: from PSNR to hybrid metrics," *IEEE Trans Broadcast.*, vol. 54, no. 3, pp. 660-668, Sep. 2008.

권순찬 (Soon-chan, Kwon)



2012년 2월 광운대학교 전자공학과 졸업
 2012년 3월~현재 광운대학교 전자공학과 석사
 <관심분야> super resolution, 영상복원

유지상 (Jisang Yoo)



1985년 2월 서울대학교 전자공학과 학사
 1987년 2월 서울대학교 전자공학과 석사
 1993년 5월 Purdue Univ. EE, ph.D.
 1997년 9월~현재 광운대학교 전자공학과 교수

<관심분야> 웨이블릿 기반 영상처리, 영상압축, 영상인식, 비선형 신호처리