

## 하이브리드 업샘플링을 이용한 베이시안 초해상도 영상처리

박종현, 강운기

연세대학교 전기전자 공학부 TMS 사업단

### Super-Resolution Image Processing Algorithm Using Hybrid Up-sampling

Jong Hyun Park, Moon Gi Kang

Yonsei University Institute of TMS Information Technology

**Abstract** - 본 논문에서는 베이시안 초해상도 영상처리시 저해상도 영상들을 고해상도 격자에 맞게 정합해서 업샘플링(upsampling)을 하는 새로운 방식에 대해 제안한다. 제안하는 업샘플링 방식은 각 장을 따로 보간하는 방식과 달리 여러 저해상도 영상의 고주파 정보가 고해상도 영상 격자의 모든 위치에 적절히 영향을 미칠 수 있도록 여러 장의 저해상도 영상의 고주파 정보를 함께 사용하여 보간한다. 보간하는 방법은 B-스플라인(B-Spline) 기반 비정규 리샘플링(non-uniform resampling)을 기반으로 초해상도 영상처리에 맞도록 적용한다. 실험결과를 통해 일반적으로 적용되는 0-삽입(zero-padding) 업샘플링 방식과 쌍일차 보간법(bilinear interpolation) 등을 적용할 때의 효과를 살펴보고, 제안하는 방식이 일반적인 방식을 사용하는 것에 비해 정량적, 정성적으로 고해상도 정보를 더 정확히 생성해내는 것을 확인한다.

#### 1. 서 론

디지털 및 아날로그 영상 입력 시스템을 통해 촬영되는 모든 영상들은 광학계와 센서의 한계에 의해 흐림 현상, 잡음, 해상도 저하 등 여러 가지 열화과정을 거쳐 저장된다. 그 중 영상신호의 해상도를 향상하기 위해서는 우선 영상 입력시스템의 광학계와 센서 자체의 해상력을 높여야 한다. 그러나 이러한 하드웨어적인 성능 향상을 위해서는 매우 높은 가격향상을 감수해야 한다. 또한 센서의 해상력을 높이기 위해 단위면적 당 화소 수를 줄이게 되면, 화소 당 센서 면적이 상대적으로 좁아지기 때문에, 결과적으로 영상이 잡음에 취약하게 되는 단점이 생긴다. 이러한 하드웨어적인 해결방식 대신 신호처리 이론을 사용한 초해상도 영상처리라는 저가의 저해상도 입력장치를 이용해 서로 상대적 부화소 단위의 움직임 차이를 가지는 여러 장의 저해상도 영상들을 촬영하고, 이들이 가지는 상호 보완적인 정보들을 소프트웨어적으로 종합하여 한 장의 고해상도 영상으로 융합하는 것이 가능하다 [1]-[3].

본 논문에서는 여러 가지 초해상도 영상처리 알고리즘 중, 베이시안 초해상도 영상처리 알고리즘을 기반으로 한 새로운 초해상도 영상처리 방법을 제안한다. 일반적으로 베이시안 초해상도 영상처리 과정에서 저해상도 영상들에 담긴 정보의 한계로 인한 화질의 저하는 영상에 대한 사전지식을 활용하는 정규화 제한조건의 비중을 키움으로서 극복한다. 저해상도 영상에 담긴 정보가 일부 영역적으로 치우치거나 저해상도 영상의 수가 충분치 못할 경우, 융합된 영상에는 부정적인 에러가 생기게 되는데, 이러한 현상의 근본적인 원인은 특정 위치에 고주파 정보가 부족함에도 불구하고 각 채널을 고해상도 격자위에 업샘플링하여 업데이트할 때 각 채널별로 독립적으로 진행하게 되므로 정보가 없는 부분은 계속해서 없는 채로 보정이 이루어지지 않기 때문에 발생한다. 본 논문은 따라서 이러한 기존의 방식 대신 관찰영상들로부터 얻은 고주파 정보를 모든 채널을 함께 참조해서 통합적으로 융합하여 고해상도 격자위에 없도록 하는 하이브리드 업샘플링 방식을 사용할 것을 제안한다.

논문의 구조는 다음과 같다. 2장에서 베이시안 초해상도 영상처리에 대해 살펴보고, 제안된 하이브리드 업샘플링 기반 초해상도 영상처리 알고리즘을 알아본다. 3장에서 실험결과를 통해 본 알고리즘의 우수성을 검증하고, 4장에서 논문을 마무리한다.

#### 2. 본 론

##### 2.1 베이시안 초해상도 영상처리

초해상도 영상처리는 열화된 영상을 복원하는 영상복원의 연장선상에서 접근되었다. 열화되지 않은 이상적인 영상을  $x$  라 하고, 열화된 영상을  $y$  라고 하자. 이 때  $x, y$  는 각각 2차원 영상을 사전기록순(lexicographically)으로 1차원 벡터로 정리한 것이다. 영상  $x$  가 흐림현상과 더해진 잡음에 의해 열화되었다고 할 때, 열화 영상  $y$  의 형성은

$$y = Hx + n \quad (1)$$

과 같이 모델화될 수 있다. 여기서  $H$  는 흐림현상을 나타내는 행렬이며,  $n$  은 더해진 잡음이다. 베이시안 확률 모델에 의한 추정기법, 특히 최대사후확률(Maximum a posteriori, MAP) 방법에서는 주어진 관찰된 신호  $y$  를 기반으로, 원 신호  $x$  의 사후 확률밀도함수(posterior probability density function, posterior PDF)를 최대화하는 신호를 찾음으로써 열화 전의 원신

호의 추정값  $\hat{x}$  을 추정한다. 즉,

$$\hat{x} = \operatorname{argmax}_x p(x|y) \quad (2)$$

이다. 여기서  $p(x|y)$  는  $x$  의 사후확률이다. 이 때,

$$p(x|y) = \frac{p(y|x)p(x)}{p(y)} \quad (3)$$

인데,  $p(y)$  는  $x$  에 상관없는 함수이므로

$$\begin{aligned} \hat{x} &= \operatorname{argmax}_x p(y|x)p(x) \\ &= \operatorname{argmax}_x \{ \ln p(y|x) + \ln p(x) \} \end{aligned} \quad (4)$$

가 된다. 그러므로, MAP 방법은 사전확률밀도함수(prior PDF)  $p(x)$  에 담긴 원 신호의 사전정보와 가능성함수(likelihood PDF)에 담긴 관찰된 신호  $y$  에 대한 정보에 가장 확률적으로 가까운 값을 추정함으로써 원 신호의 추정값을 찾는다. 영상의 형성 모델을 기초로 한 가능성 함수는 영상에 더해진 잡음이 화소간에 서로 독립적인 가우시안 확률분포를 갖는다고 가정할 때,

$$p(y|x) = \frac{1}{N} e^{-\frac{\|y - Hx\|^2}{2\sigma_n^2}} \quad (5)$$

과 같이 나타낼 수 있다. 이 때  $N$  은 확률분포함수의 적분값이 1이 되기 위한 상수이며,  $\sigma_n^2$  은 가우시안 잡음의 분산이다. 한편, 원 영상의 사전정보는 영상의 고주파 정보가 가우시안 분포를 따른다고 하는 기본적인 가정에 근거할 때,

$$p(x) = \frac{1}{M} e^{-\frac{\|Cx\|^2}{2\sigma_c^2}} \quad (6)$$

와 같이 나타낼 수 있다. 역시 이 때  $M$  은 확률분포함수의 적분값 1로 하기 위한 상수이며,  $C$  는 라플라시안(Laplacian) 고주파 통과 필터이고,  $\sigma_c^2$  는  $x$  의 고주파 정보의 분포 정도를 나타내는 분산값이다. 수식 (5)와 (6)의 두 가지 확률분포를 수식 (4)의 MAP에 적용하여 정리하면,

$$F(x) = \|y - Hx\|^2 + \frac{\sigma_n^2}{\sigma_c^2} \|Cx\|^2 \quad (7)$$

과 같은 함수를 최소화하는  $x$  를 추정하는 형태로 나타낼 수 있다. 이 형태는 제한된 최소 자승법(constrained least squares, CLS) 방식으로 해석될 수 있다. 수식 (7)에서 앞의 부분은 관찰모델을 기반으로 한 추정대상 영상과 관찰된 영상 사이의 에러라고 생각할 수 있으며, 이 때 두 번째 항은 영상의 고주파 정보의 양을 제한조건으로 추가한 형태가 된다. CLS에서 제한 조건의 반영정도를 조절하는 CLS의 정규화 파라미터(regularization parameter)는 관찰모델의 충실도를 나타내는 앞쪽 항과 사전정보에 대한 충실도를 나타내는 뒤쪽 항의 반영 정도를 조절하는데,  $\sigma_n^2/\sigma_c^2$  가 이와 같은 역할을 하고 있다고 해석할 수 있다.

한 장에 대해 적용하는 베이시안 영상복원을 여러 장의 저해상도 영상을 열화된 관찰 영상으로 하고 한 장의 고해상도 영상을 복원 대상인 원 영상으로 하는 초해상도 영상처리에 적용할 수 있다. 고해상도 영상이 움직임과 흐림현상, 다운샘플링(downsampling), 잡음 등 여러 단계에 의해 열화되었다고 가정할 때 영상 형성에 대한 모델은

$$\begin{aligned} y_k &= W_k x + n_k \\ &= DB_k M_k x + n_k, \quad \text{for } k=1,2,\dots,p, \end{aligned} \quad (8)$$

과 같다. 여기서  $p$  는 저해상도 영상의 개수,  $W_k$  는  $k$  번째 영상에 대한 열화 행렬이며, 각각 다운샘플링, 흐림, 움직임을 나타내는 행렬  $D, B_k$  와  $M_k$  의 조합으로 나뉘질 수 있다.  $n_k$  는 평균이 0인 더해진 가우시안 잡음이다. 수식 (8)의 관찰 모델을 기반으로 한 초해상도 영상처리는

$$F(x) = \sum_{k=1}^p \|y_k - W_k x\|^2 + \gamma \|Cx\|^2 \quad (9)$$

와 같은 함수를 최소화하는 해를 구함으로써 고해상도 영상을 얻게 된다. 여기서  $\gamma$  는 정규화 파라미터이다. 수식 (9)의 우변의 첫 번째 항  $\sum_{k=1}^p \|y_k - W_k x\|^2$  은 관찰된 영상에 대한 해 영상의 충실도를 의미하며, 두 번째 항  $\gamma \|Cx\|^2$  은 해의 고주파 정보로 해 영상의 부드러운 정도를 반영한다.  $\gamma$  의 값의 변화에 따라 관찰된 영상과의 충실도를 어느 정도 손해 보면서 해 영상을 정규화 할지가 결정된다.

## 2.2 하이브리드 업샘플링

반복적 그래디언트 하강 방식(iterative gradient descent method)을 통해 수식 (9)를 최소화하는 해를 얻을 수 있다. n번 반복 시 해의 추정치  $\hat{x}^n$ 를 사용한 n+1번째의 반복해  $\hat{x}^{n+1}$ 은

$$\hat{x}^{n+1} = \hat{x}^n - \beta^n \nabla_x F(x) \Big|_{x=\hat{x}^n} \text{ for } n=0,1,2,\dots \quad (10)$$

이다. 여기서  $\beta^n$ 은 n번째 반복에서 해를 갱신하는 정도를 조절하는 파라미터이다. 반복해의 초기값  $\hat{x}^0$ 는 일반적으로 저해상도 영상들 중의 하나를 고해상도 영상 크기로 보간한 영상을 사용한다. 수식 (10)의  $\nabla_x F(x)$ 는

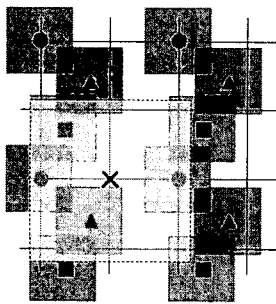
$$\nabla_x F(x) = \sum_{k=1}^2 2W_k^T(W_k x - y_k) + 2\gamma C^T C x \quad (11)$$

이다. 이 반복적 그래디언트 하강방식의 해를 구하는 과정을 해석해보자  $W_k x - y_k$ 는  $\hat{x}^n$ 과  $y_k$ 사이의 저해상도 격자 위치에서의 관찰모델을 기반으로 한 차이, 즉 n번째 추정치에 아직 들어있지 않은  $y_k$ 의 고주파 정보를 계산한 것으로 볼 수 있다.  $W_k^T$ 는 열화행렬  $W_k$ 의 전치행렬로, 풀어서 쓰면  $M_k^T B_k^T D^T$ 와 같다.  $M_k^T$ ,  $B_k^T$ 와  $D^T$ 는 각각 고해상도 격자 상의 움직임 보정, 고해상도 격자 상의 흐름 효과, 그리고 저해상도 영상을 고해상도 격자 위치로 업샘플링하는 역할을 한다. 즉,  $y_k$ 에 남아있는 새로운 정보는  $W_k^T$ 에 의해서 고해상도 영상에 정합되어  $\hat{x}^n$ 에 더해져서 새로운 추정값  $\hat{x}^{n+1}$ 을 생성하게 되는 것이다. 이때  $\gamma$  값으로 상대적으로 큰 값을 사용하면, 영상의 고주파 정보를 더 최소화시키게 되므로 결과 영상이 더 부드러워지게 된다. 그런데, 원래  $D$ 의 의미를 그대로 가진  $D^T$ 를 쓰게 되면,  $D^T$ 는 업샘플링으로 0-삽입 보간을 수행하게 되므로, 다수의 저해상도 영상이  $D^T$ 를 포함한  $W_k^T$ 에 의해 정합이 되고 나서도 모든 고해상도 격자 위치를 다 포괄하지 못하면 어느 위치에는  $y_k$ 의 정보가 반복해를 구하는 과정에서 계속해서 업데이트 되지만, 어떤 위치에는 정보가 부족해서 계속해서 업데이트되지 않는 부작용이 생기게 된다. 이것은 저해상도 영상의 수가 많다고 하더라도 움직임이 한쪽 위치로만 쏠려 있을 때에도 발생하며, 영상의 수가 적으면 더욱 심화된다.

이러한 현상은 뒷부분에 있는 정규화 부분의 역할을 더 키워서, 즉  $\gamma$ 의 값을 크게 해서 부드러운 영상을 추구함으로써 해결한다. 그러나 정규화 부분을 키우는 대신, 업샘플링 합수를 바꿈으로써 이러한 문제를 해결할 수 있다. 본 논문에서는 단순히 각 채널을 0-삽입 업샘플링하는 대신, 각 저해상도 채널의 정보를 함께 참조해서 업샘플링을 수행한다. 본 논문에서는 이와 같은 업샘플링 방식을 하이브리드(hybrid) 업샘플링이라 지칭한다. 제안하는 업샘플러를 포함한 변형된  $\nabla_x F(x)$ 는  $2Id + 2\gamma C^T C x$ 와 같다. 여기서

$$d = \begin{bmatrix} W_1^T(W_1 x - y_1) \\ W_2^T(W_2 x - y_2) \\ \vdots \\ W_p^T(W_p x - y_p) \end{bmatrix} \quad (12)$$

이고,  $I$ 는 여러장의 영상의 보편적인 정보가 함께 담겨있는  $d$ 를 기반으로 최종적으로 고해상도 격자에 보간하는 행렬이다. 이를 그림으로 나타낸 것이 그림 1이다.



〈그림 1〉 고해상도 격자 상의 저해상도 영상들과 보간 과정의 개념도

그림 1에서 원, 삼각형, 사각형 등의 형태로 표시된 것들은 고해상도 격자에 정합된 저해상도 영상 화소를 의미하며, x표시는 현재 업데이트해야 하는 고해상도 격자 위치이다. x 주변의 점선 영역은 현재 위치에서 참조할 저해상도 화소들의 위치를 나타내고 있다. 따라서 I는 점선 안쪽 영역에 포함된 화소들을 바탕으로 통합적으로 보간을 수행하게 된다. 본 논문에서는 보간 방법으로 B-스플라인 방식을 사용한다.

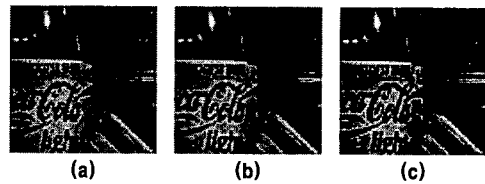
## 3. 실험 결과

본 알고리즘을 검증하기 위해 고해상도 영상을 영상모델을 이용해서 여러 장의 부화소 단위 움직임이 있는 저해상도 영상으로 다운샘플링한 뒤, 다시 초해상도 영상처리를 이용해 고해상도 영상으로 재구성한 후 비교하였다. 제안하는 하이브리드 업샘플러 이용 알고리즘과의 비교를 위해 0-삽

입 업샘플러와 쌍일차 보간 업샘플러를 사용하였다. 실험 영상으로는 가로, 세로 방향 예지가 풍부한 512x512 크기의 "shop"영상을 원 고해상도 영상으로 가정하고, 영상의 흐름합수로 5x5 크기의 분산 0.5인 가우시안 흐름합수를 사용하였다. 잡음은 첨가하지 않았다. 첫 번째 실험을 위해 4장의 저해상도 영상을 생성하였다. 저해상도 영상 생성 시 (x, y)를 x방향, y방향의 평행이동이라고 할 때, 고해상도 영상을 고해상도 격자 상에서 각각 (0,0), (1,0), (0,1), (1,1)만큼 평행이동 하도록 하고, 가로 세로 길이가 0.5배인 영상으로 다운샘플링했다. 이 형태는 각 고해상도 격자 위치에 각각의 저해상도 영상을 정합할 경우 모든 위치에 적절한 정보가 삽입될 수 있는 이상적인 경우이다. 두 번째 실험 영상은 위와 동일한 영상에서 저해상도 영상을 (0,0), (1,0), (0,1) 세 영상만 사용한 경우이다. 이때는 영상의 수가 적게 되므로, 고해상도 재구성된 영상에 에러가 생길 확률이 그만큼 커진다. 두 실험 모두 저해상도 영상을 다시 가로 세로 2배의 크기로 확대하는 실험으로 진행하였다. 본 실험의 수치적 비교는 최고치 신호 대 잡음비(peak signal to noise ratio, PSNR)를 사용하였다. 실험 결과 영상을 각각 원 "shop"영상과 비교하여 PSNR수치를 얻은 결과가 표 1과 같다.

〈표 1〉 "shop"영상의 PSNR 비교

	0-삽입 방법	쌍일차 보간법	하이브리드 보간법
실험 1	44.56	34.74	45.91
실험 2	30.19	32.29	35.17



〈그림 2〉 "Shop"영상의 실험 2 영상 set을 사용한 재구성 결과 (a) 0-삽입 방법, (b) 쌍일차 보간법, (c)하이브리드 보간법

표 1에서 볼 수 있는 바와 같이, 일반적으로 사용하는 0-삽입 방법은 필요한 정보가 잘 갖춰져 있는 경우에는 영상을 정확히 복원해내지만, 영상의 수가 적은 경우에는 영상을 제대로 복원해내지 못하는 것을 알 수 있다. 한편 쌍일차 보간법을 사용한 경우는 반대로 전반적으로 고른 결과를 내지만, 어떤 경우에도 정확한 결과를 내지는 못한다. 그러나 하이브리드 보간법을 사용한 재구성 결과는 영상의 정보가 충분한 곳에서도 가장 좋은 결과를 내었으며, 영상의 수가 적은 경우에도 0-삽입 방법에 비해 약 5dB, 쌍일차 보간법에 비해 약 3dB의 성능 향상을 보였다. 그림 2는 실험 2에 대한 결과 영상들을 일부분 잘라내어 삽입한 것이다. 그림 2의 화질을 정성적으로 비교해보아도 그림 2-(c)의 하이브리드 보간법 결과영상이 가장 좋은 화질로 원 영상의 정보를 재구성해내고 있는 것을 확인할 수 있다.

## 4. 결 론

본 논문에서는 하이브리드 업샘플링을 이용한 베이스인 초해상도 영상 재구성 방법을 제안하였다. 제안된 알고리즘은 변형된 업샘플링 방식을 사용하여, 각 저해상도 채널의 정보를 골고루 이용해서 각 고해상도 격자 위치에 고주파 정보를 채워 넣기 때문에 더 정확한 고해상도 영상 추정이 가능하게 되었다. 실험결과를 통해서도 제안된 알고리즘이 고해상도 정보를 더 정확히 추정해낼 수 있는 것을 수치적으로, 정량적으로 평가할 수 있었다.

## Acknowledgement

이 논문은 2007년도 교육인적자원부 BK21 사업의 일환인 연세대학교 전기 전자 공학부 TMS 사업단과 서울시 산학연협력사업으로 구축된 서울 미래형콘텐츠컨버전스 클러스터 공동지원으로 연구되었음

## [참 고 문 헌]

- [1] S. Park and M. G. Kang "Superresolution Image Reconstruction -A Technical Overview," IEEE Signal Processing Magazine, Vol. 20, no. 3, pp. 21-36, May 2003.
- [2] R. Y. Tsai and T. S. Huang, "Multipleframe image restoration and registration," Advances in Computer Vision and Image Processing, vol. 1, pp. 317-339, JAI Press Inc., 1984.
- [3] R. C. Hardie, K. J. Barnard, J. G. Bognar, E. E. Armstrong, and E. A. Watson, "High-resolution image reconstruction from a sequence of rotated and translated frames and its application to an infrared imaging system," SPIE Optical Engineering, vol. 37, no. 1, pp. 247-260, Jan. 1998.