

부화소 단위의 움직임 정보를 고려한 순차주사화

박민규, 강문기

연세대학교 전기 전자공학과

Sub-pixel motion compensated deinterlacing technique

Min Kyu Park, Moon Gi Kang

Department of Electrical and Electronic Yonsei University

E-mail : redsky@yonsei.ac.kr

요 약

HDTV(high-definition television)와 개인용 컴퓨터의 발전에 따라 비월주사(interlaced scanning) 방식의 신호와 순차주사(progressive scanning) 방식의 신호 상호간의 변환에 대한 요구가 점점 늘어나고 있다. 이에 따라 비월주사 방식을 순차주사 방식으로 바꾸어주는 순차주사화(deinterlacing)가 활발히 연구되고 있다. 본 논문에서는 부화소(sub-pixel) 단위의 움직임 정보를 고려한 영상재결합 방법을 기반으로 한 순차주사화 알고리즘을 제안한다. 그러나 국부 움직임이 있는 영역에서는 영상재결합이 효과적이지 않으므로 적절한 문턱값을 이용하여 이러한 영역을 결정하고 에지를 이용한 보간 방법으로 순차주사화 한다. 또한 잘못된 움직임 추정으로 야기되는 문제들을 해결하기 위해, 각 영역과 입력 영상에 따라 다른 움직임 에러를 고려해서 정규화 값(regularization parameter)을 결정함으로써 움직임 에러를 효율적으로 제거하였다. 제안된 방법은 주관적 측면과 객관적인 측면에서 모두 우수한 결과를 실험적으로 보였다.

1. 서 론

최초의 텔레비전은 대역폭의 제한과 메모리의 제약으로 인해 비월주사 방식을 채택하게 되었다. 그러나 비월주사 방식의 영상을 고해상도의 영상으로 바꾸기 위해서는 비월주사 방식의 필드(field)를 순차주사 방식의 프레임(frame)으로 바꾸어 주는 순차주사화 선형되어야 한다. 이러한 순차주사화는 크게 하나의

필드를 사용하여 프레임을 만드는 필드내 보간과 여러 필드를 사용하는 필드간 보간으로 나눌 수 있다. 필드내 보간은 처리 시간과 구현이 간단한 선형 보간과 경계의 방향을 고려한 대각 보간(edge-adaptive interpolation)[1], 가중치를 사용한 미디언 필터(median filter)[2], 그리고 미디언 필터와 저역통과필터의 결합 형태로 사용하는 방법(NMF/MAF: nested median filter/median averaging filter)등이 있다[3]. 이는 한 필드내에서 연산이 이루어지므로 계산량이 적어 연산 속도가 빠르다는 장점이 있으나, 원영상에 비해 고주파 성분의 에너지가 많이 손실되는 단점이 있다. 이에 반해 필드간의 보간은 주로 움직임을 고려하여 보정하는 방법과[4-5], 주파수 영역에서 부화소 단위의 움직임을 고려한 고해상도 영상 재구성 방법이 있다[6]. 이러한 방법은 한 필드에 의한 방법보다 고해상도를 얻을 수 있으나 계산량이 많으며, 필드간의 움직임이 일정한 속도로 변하지 않는 경우와 필드간 물체의 모양이 변화하는 경우에 대해서는 좋은 결과를 기대하기 어렵다는 단점을 가지고 있다. 그래서 필드내 보간과 필드간 보간 각각의 장점을 살리고자 한 것이 필드내 보간과 필드간 보간을 혼합하여 사용한 방법이다[7][8].

본 논문에서는 국부움직임이 있는 영역과 그렇지 않은 영역을 공간영역의 통계적 특성을 고려한 문턱값을 이용하여 구분하였다. 이렇게 구분된 영역중에서 국부 움직임이 두드러진 영역은 가중치를 사용한 대각 보간으로 순차주사화를 하였고, 전역움직임이 두드러진 영역에서는 움직임 추정의 에러를 고려하여 각 채널마다 다른 정규화값을 사용하여 움직임 에러에 강하게 순차주사화 하였다.

2. 제안된 순차주사화

주어진 시간 t 에 대해서 (x,y) 의 좌표에서의 밝기의 값을 $I(x,y,t)$ 라 하자. 필드내 보간에 의해서 복원된 값을 $I_S(x,y,t)$ 라하고, 필드간 보간에 의해서 복원된 값을 $I_T(x,y,t)$ 라 하면 복원된 값은 일반적으로 다음과 같이 식(1)으로 정리된다[8].

$$I(x,y,t) = \epsilon I_S(x,y,t) + (1-\epsilon) I_T(x,y,t) \quad (1)$$

여기에서 ϵ 는 0과 1사이의 적당한 값으로 여러 가지 방법[7][8]에 의해 결정되어 진다. 제안된 방법은 식(1)을 기본으로 하여 다음과 같이 나타낼 수 있다.

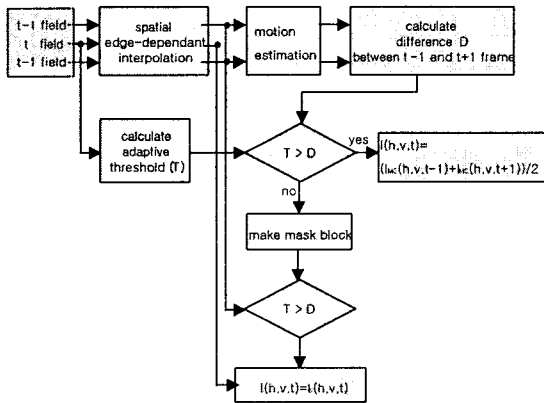


그림 1 제안된 순차주사화 알고리즘 블록도

2.1 필드내 보간

일반적으로 필드내 보간 중에 경계의 방향을 고려한 대각 보간이 가장 성능이 가장 우수하다. 그러나 이러한 방법은 에지를 잘못 찾을 경우 오히려 물체의 경계를 제대로 보존하지 못한다는 단점을 갖기 때문에 이를 극복하기 위해서, 벡터를 이용한 가중치 대각 보간을 사용한다. 우선, 벡터 $\mathbf{u}(n)$ 와 $\mathbf{v}(m)$ 를 각각

$$\mathbf{u}(n) = \begin{bmatrix} I(x+n-1, y-1, t) \\ I(x+n, y-1, t) \\ I(x+n+1, y-1, t) \end{bmatrix}, \quad (2)$$

$$\mathbf{v}(m) = \begin{bmatrix} I(x+m-1, y+1, t) \\ I(x+m, y+1, t) \\ I(x+m+1, y+1, t) \end{bmatrix} \quad (3)$$

라고 하면 WAD(weight absolute difference)는

$$D(n,m) = |\mathbf{u}(n) - \mathbf{v}(m)|_c \quad (4)$$

이고, c 일반화 상수(normalized weight)이다. 이때, (n',m') 는

$$(n',m') = \min \arg D(n,m) \quad (5)$$

을 통해 얻을 수 있다. 얻어진 (n',m') 을 통해 우리가 복원하려는 $I(x,y,t)$ 는 다음과 같은 식(6)을 통해 구할 수 있다.

$$I(x,y,t) = \frac{I(x+a, y-1, t) + I(x+b, y-1, t)}{4} + \frac{I(x+c, y+1, t) + I(x+d, y+1, t)}{4} \quad (6)$$

여기에서

$(a,b,c,d) =$

$$\begin{cases} (0,0,0,0) & \text{if } (n',m') = (-1,-1), (0,0), (1,1) \\ (-1,0,0,1) & \text{if } (n',m') = (-1,0), (0,1) \\ (-1,-1,1,1) & \text{if } (n',m') = (-1,1) \\ (0,1,-1,0) & \text{if } (n',m') = (0,-1), (1,0) \\ (1,1,-1,-1) & \text{if } (n',m') = (1,-1) \end{cases} \quad (7)$$

2.2 움직임 추정

전역 움직임 추정의 정확도를 높이기 위해 움직임 추정은 패리티(parity)가 같은 프레임을 사용하고 BMA(block matching algorithm)를 통해 화소 단위의 움직임 벡터를 구한다. dx 와 dy 를 시간 $t-1$ 과 $t+1$ 사이의 영상의 움직임 벡터라 하면, $I(x+dx/2, y+dy/2, t-1)$ 와 $I(x-dx/2, y-dy/2, t+1)$ 의 차의 절대값이 그림 2이다. 이때, 국부적인 움직임이 두드러진 경우와 물체의 모양이 변하는 부분에서 밝기의 차가 커지게 된다.

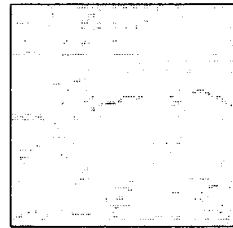


그림 2. 두 프레임의 차 영상

일반적으로, 공간영역에서 에지와 같이 밝기의 분산이 큰 영역에서는 필드내 보간이 필드간 보간보다 우수한 성능을 갖는다. 이러한 특성을 통해 문턱값(T)을 실험적으로 정한다.

$$T = \frac{2}{\pi} \times T_{\max} \times \tanh^{-1} \left(\frac{\text{Var}}{k} \right) \quad (8)$$

여기에서 Var 은 공간 영역에서의 밝기의 분산이고 T_{\max} 와 k 는 상수이다.

그림 2에서 밝기가 문턱값 이상인 부분은 그림 3과 같이 마스크(mask)들로 분리하여 2.1에서의 방법으로 필드내 보간을 한다.

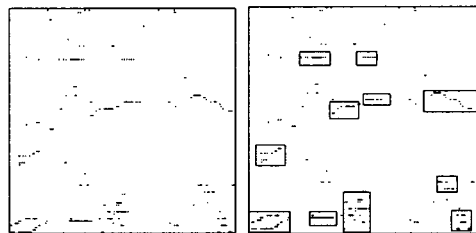


그림 3 차 영상과 마스크

2.3 영상재결합 방법을 기반으로 한 순차주사화

고해상도 영상이 $L_1N_1 \times L_2N_2$ 크기를 가지고 있다고 하고, 이를 사전배열식(lexicographical ordering)으로 나타낸 벡터를 $\mathbf{x}=[x_1, x_2, \dots, x_N]^T$, $N=L_1N_1 \times L_2N_2$ 라 하자. 이때, 저해상도의 영상들은 $N_1 \times N_2$ 의 크기를 가지고 있다. 이때 k 번째 저해상도의 영상을 사전배열식으로 나타낸 벡터는 $\mathbf{y}_k=[y_{k,1}, y_{k,2}, \dots, y_{k,N}]^T$, $k=1, 2, \dots, p$ 이고 $M=N_1 \times N_2$ 이다.

저해상 영상의 화소는 고해상도의 영상의 화소들을 가중치를 주어서 합한 형태로 다음과 같이 모델링 할 수 있다[9].

$$y_m = \sum_{r=1}^N w_{m,r}(s)x_r + \eta_m, \quad m=1, 2, \dots, pM \quad (9)$$

여기에서 가중치 $w_{m,r}(s)$ 는 k 번째 저해상도 영상에 있는 m 번째 화소에 대한 r 번째 고해상도 영상의 기여도이다. 이때 s 는 수직, 수평 방향으로의 부화소 단위의 움직임 정보를 의미한다. 식(9)를 행렬로 표현하면,

$$\mathbf{y} = \mathbf{W}_s \mathbf{x} + \mathbf{n} \quad (10)$$

이다.

주어진 저해상도 영상(필드) \mathbf{y} 와 부화소 단위로 추정된 움직임 정보로부터 고해상도 영상(프레임) \mathbf{x} 를 복원하기 위해 다음과 같은 식을 세울 수 있다.

$$\hat{\mathbf{x}} = \arg \min M(\mathbf{x}) \quad (11)$$

여기에서

$$M(\mathbf{x}) = \frac{1}{2} \sum_{p=1}^P \lambda_p(\mathbf{x}) \sum_{m=1}^M (y_{p,m} - \sum_{r=1}^N w_{p,m,r}(\mathbf{x})x_r)^2 + \frac{1}{2} \sum_{i=1}^N (\sum_{j=1}^N C_{i,j}x_j)^2 \quad (12)$$

이고, $\lambda_p(\mathbf{x})$ 는 정규화 파라미터이고 C 는 고역통과여파기(high pass filter)이다. 좌변은 주어진 저해상도 영상과 고해상도 영상으로 예측된 저해상도 영상 사이의 차이를 줄여주는 부분이고, 우변은 추정 잡음을 보정해주는 부분이다. 이 때 각 채널의 움직임 에러를 고려하기 위해서는 $\lambda_p(\mathbf{x})$ 가 식(12)에서 우변에 붙어야 한다. 만약 우변에 $\lambda_p(\mathbf{x})$ 가 붙게 되면 움직임 에러가 하나라도 큰 경우에는 식에 의해서 모든 채널의 영상에 대해 지나친 블러(blur)가 된다. 이때 $\lambda_p(\mathbf{x})$ 는 각 채널에 움직임에러와 고주파 에너지의 비례관계로 나타낼 수 있다.

$$\lambda_p(\mathbf{x}) = \frac{\|C\mathbf{x}_p\|^2}{\|y_p - \mathbf{W}_p\mathbf{x}\|^2 + \delta_p} \quad (13)$$

식(11)은 미분 강하(Gradient descent)로 해를 구할 수 있다.

$$\mathbf{x}_k^{n+1} = \mathbf{x}_k^n - \epsilon^n \mathbf{g}_k(\mathbf{x}_k^n) \quad (14)$$

여기에서

$$\mathbf{g}_k(\mathbf{x}_k^n) = \sum_{p=1}^P \lambda_p w_{p,m,k}(\mathbf{x}) \sum_{m=1}^M (y_{p,m} - \sum_{r=1}^N w_{p,m,r}(\mathbf{x})x_r) + \sum_{i=1}^N C_{i,k} \sum_{j=1}^N C_{i,j}x_j \quad (15)$$

이고 ϵ^n 은 $M(\mathbf{x}_k^{n+1}) = M(\mathbf{x}_k^n - \epsilon^n \mathbf{g}_k(\mathbf{x}_k^n))$ 을 ϵ^n 으로 미분하여 값을 얻는다.

3. 실험 및 결과

제안된 알고리즘의 성능을 비교하기 위해 실험에는 "salesman" 영상이 사용되었다. 우선 영상을 사용하여 세 영상에 각각 (-2.25, -4.25), (0, 0), (2.0, 4.25)의 움직임을 주고 가운데로부터 256×256 의 영상을 실험에 사용하였다. 그림 4는 원 영상이고 그림 5는 128×256 의 비월주사된 영상들이다.

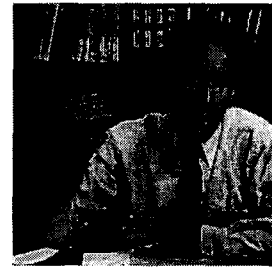


그림 4. 원 영상

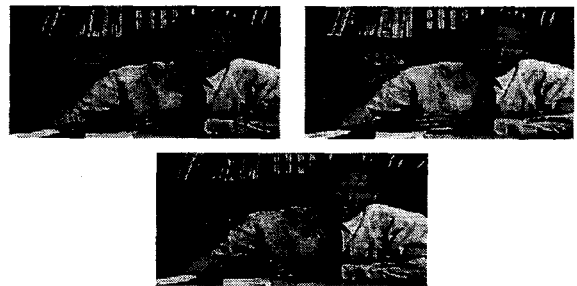


그림 5. 비월주사된 영상들

그림 6은 선형 보간에 의한 결과이고, 그림 7은 BMA에 의한 움직임 보정에 의한 결과이다. 선형 보간에 경우 전체적인 선명도가 제안된 방법이나 BMA에 비해 현저하게 떨어지는 것을 알 수 있다.

선형보간은 사선 경계에 경우에 그 성능이 현저하게 떨어지게 되는데, 좌측 하단에 보면 계단 모양과 같이 좋지 못한 현상이 생긴다. BMA의 경우에는 quad-tree 방법을 사용하여 영역을 분할해 가면서 움직임 보정을 했다. 움직임 보정이 계속해서 안되는 경우에 대해서는 2.1에서의 방법으로 대각 보간을 했다. 이러한 방법은 필드내 방법보다는 우수한 성능을 갖는다. 그러나 손가락에서와 같이 움직임이 정확하게 찾아지지 않는 경우에 대해서는 좋지 못한 결과를 낳는 것을 볼 수 있다. BMA의 경우에는 부화소 단위의 전역 움직임을

고려하지 못함으로써 미세하게나마 전체적으로 제안된 알고리즘의 결과보다 다소 성능이 떨어지는 것을 알 수 있다. 이를 수치적으로 비교하기 위해 PSNR(peak signal-to-noise ratio)을 사용했다. PSNR의 식(15)와 같다.

$$PSNR=10 \log_{10} \frac{255^2}{H \times V} \sum_{h=0}^{H-1} \sum_{v=0}^{V-1} (I(h, v, t) - \hat{I}(h, v, t))^2 \quad (15)$$

여기에서 H는 영상의 가로방향의 크기를 V는 영상의 세로 방향의 크기를 나타내고 $I(h, v, t)$ 는 원 영상의 밝기 값을 $\hat{I}(x, y, t)$ 는 추정된 영상의 밝기를 의미한다.

algorithm	PSNR(dB)
linear	25.34
BMA	28.31
Proposed	30.60

표 1 제안된 방법과의 비교



그림 6. 선형보간

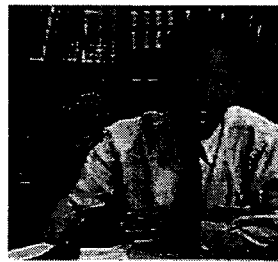


그림 7 BMA의 결과



그림 8. 제안된 방법

4. 결론

본 논문은 필드에 대해서 시간과 공간적인 영역에서 순차주사화 알고리즘을 제안한다. 특히 부화소 단위의 움직임 정보를 사용하였으며, 부정확한 움직임 정보에 대해서도 정규화 파라미터를 각 채널에 따라 달리 줌으로써 각 채널에 따라 움직임 에러를 보정할 수 있었다. 또한 실험을 통해서 제안된 방법이 기존의 다른 순차주사화 방법보다 향상된 성능을 갖는다는 것을 실험을 통해 확인할 수 있었다.

참고 문헌

- [1] Yeong-Taeg Kim, "Deinterlacing Algorithm Based on Sparse Wide Vector Correlations", SPIE vol 2727, Pages 89 - 99. 1996
- [2] Janne Juhola, Ari Nieminen, Juha Salo and Yrjö Neuvo, "Scan Rate Conversion Using Weighted Median Filtering", IEEE, ISCAS, 1996
- [3] H. Hwang, "Interlaced to Progressive Scan Conversion For HD-MAC Application", Consumer Electronics, IEEE Transactions on Pages: 151 -156 Aug. 1992 Vol.38
- [4] Wang, F.M. , Anastassiou, D. and Netravali, A.N. "Time-recursive deinterlacing for IDTV and pyramid coding", Circuits and Systems, 1990., IEEE International Symposium on , 1990 , Page(s): 1306 -1309 vol.2
- [5] Delogne, P. , Francis, P. and Vandendorpe, L. "Conversion from interlaced to progressive formats by means of motion compensation based techniques", Image Processing and its Applications, 1992., International Conference on , 1992 , Page(s): 409 -412
- [6] Ryu, C. and Kim, S.P. "Deinterlacing using motion compensated local spectra ", Signals, Systems and Computers, 1995. Conference Record of the Twenty-Ninth Asilomar Conference on Volume: 2 , 1996 , Page(s): 1394 -1397 vol.2
- [7] Kenji Sugiyama and Hiroya Nakamura, "A Method of Deinterlacing with motion compensated interpolation", Consumer Electronics, IEEE Transactions on Volume: 45 3 , Aug. 1999 , Page(s): 611 -616
- [8] Kovacevic, J. Safranek, R.J. Yeh, E.M. "Deinterlacing by Successive Approximation." Image Processing, IEEE Transactions on Volume: 6 2 , Feb. 1997 , Page(s): 339 -344
- [9] 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," Opt. Eng., vol. 37, no.1, pp.247-260, Jan.1998.