왜곡 블록 정합 방식의 기하학적 오류 보정 방법에 관한 연구

고기석, 이우섭
한신대학교 컴퓨터정보학과

Study on a Correction of Geometrically Over-Deformed Regions for Deformable Block Match Algorithm

Ki-Seok Ko, Ou-Seb Lee
Dept. of Computer and Information Science at Hanshin University

요 약

동화상 압축에서 사용되는 움직임 벡터는 영상을 일정한 크기의 블록으로 나누어 이전 영상과 가장 예측 오류가 적은 곳을 지정하여 예측한다. 블록의 이동 정도는 수직, 수평의 선형적인 움직임(translational displacements)을 가정하여 사용하기 때문에 실제 화상에서 자주 나타나는 물체의 확대 또는 축소에 의한 크기 변화, 회전, 일그러짐 등의 변화에 올바르게 예측하지 못하는 문제를 가지고 있다. 본 논문은 블록의 선형적인 이동은 물론이고 4개 노드를 자유롭게 움직임일 수 있는 왜곡(歪曲)된 블록 정합 방식의 움직임 예측 기법에 대하여 소개하고 왜곡 블록 정합 방식에서 나타날 수 있는 기하학적인 오류를 수정하는 보정 방법에 관해 논의한다.

1. 서론

동영상 압축 기법에서 중요한 한 부분을 다루고 있는 움직임 예측 기법은 기존의 표준화된 다양한 방법에서 사용되어 왔다. 예를 들어 동영상 저장용으로 주로 사용하는 MPEG-1이나, 화상회의 표준안인 H.263 등에서 디지털 동영상을 압축하기 위한 주된 방법으로 움직임 예측에 의존한다.

움직임 예측은 동화상을 구성하는 각 영상들의 시간축의 상호 연관성을 이용한 것으로, 기준 영상과 부호화할 영상간의 데이터 중복성을 찾아내서 이를 제거하는 작업이다. 동화상의 각각의 영상들은 대부분 이전 영상에서 나타났던 물체가 어느 정도 이동하여 다시 나타남으로, 물체의 이동량을 파악하여 이를 부호화하면 다음 영상을 구성할 수 있다. 이전 영상의 물체가 다음 영상에서 나타나고, 이전 프레임 영상들이 다음 프레임의 영상을 구성하는데 기반이 될 수 있다는 점을 활용한 것으로 시간축 상의 잉여 데이터(temporal redundancy)를 이용한 것이다 [1][2].

영상 압축의 표준안으로 널리 알려진 MPEG-1에서 사용하는 방법을 예로 들면, 전체 동영상을 유사한 장면을 포함한 그룹으로 나눈 후에, 각 그룹은 다시 일정간격으로 기준 영상을 정한다. 기준영상은 부호화할 영상을 구성하기 위한 예측의 기반을 구성하는 것으로 일정한 크기의 블록으로 나누어진 부호화할 영상에서 각각의 정합부분에 대한 화소 정보를 제공하는 곳이기도 하다.

블록의 정합을 찾는 과정은 오류정합함수를 최소화하는 곳을 예측함으로 이루어진다. 하지만 기존의 블록 정합 방식은 단지 수직 및 수평의 선형적인 움직임만을 가정하여 예측하기 때문에 비선형적인 움직임 영역에서는 예측 에러가 커진다. 예를 들어 카메라의 줌인(zoom in), 줌아웃(zoom out)에 의한 영상의 확대, 축소, 회전등은 선형적인 움직임이 아니며, 이곳에서는 예측된 움직임 벡터의 신뢰도가 감소한다. 하지만, 왜곡 블록정합 방식은 주어진 블록의 움직임을 수직, 수평의 선형적인 움직임만을 고려한 것이 아니라, 블록을 형성하는 네 개의 꼭지점

(node)의 자유스러운 움직임을 이용하여 정합을 찾는 방법이기 때문에 비선형 움직임에 대한 예측으로 사용된다[1]. 네 개의 꼭지점을 이용한 왜곡 정합 예측도 에러함수를 최소화하는 방법으로 정합점을 찾아가게 되며 자유스러운 꼭지점의 움직임 때문에 지나친 사각형의 왜곡이 발생하기도 한다. 예를 들어, 네 개의 꼭지점 중의 하나가 사각형의 내부로 지나치게 이동하여 삼각형을 이루거나 한 꼭지점의 내부의 각도가 180도 이상이 되는 경우가 발생할 수 있다. 이를 기하학적인 과도 변형이라고 정의하며, 계속되는 동영상의 안정적인 수치해석을 위하여 지나친 왜곡을 수정할 필요가 있다. 두 영상의 시간 차이가 매우 작은 두 프레임 사이에서는 이러한 물체의 이동이 발생하지 않는 것으로 간주하여 네 개의 꼭지점 중의 하나를 이동하여 왜곡을 수정하게 된다.

본 논문은 왜곡 블록 정합을 동영상의 움직임 예측 필요한 네 개의 모서리 이동벡터 추정 방법에 대하여 소개하고, 정형화된 블록의 왜곡과정에서 발생하는 기하학적인 과도 왜곡 (over-deformation)을 인식하는 과정과 이를 수정하는 기법을 제시한다.

2. 왜곡(歪曲) 블록 정합 방식

왜곡 블록 정합은 네 개의 모서리의 자유로운 이동을 적용하기 위하여 네 개의 모서리에서의 변위를 구하기 용이한 규격화(normalized)된 마스터 도메인(master domain)을 이용한다. 이는 왜곡 사각형의 에러 추정에 필요한 수치해석을 안정적으로 사용할 수 있게 한다[2]. 마스터 도메인은 정사각형 네 개 노드의 좌표를 1과 -1 위치시킨 규격화된 좌표계에 놓인 정사각형을 일컫는다. 이를 이용하면 순방향(forward)과 역방향(inverse) 변환에 필요한 역행렬을 유도할 필요가 없이 직접 순방향의 이동 변위를 이용하여 표현할 수가 있다. 더불어 다항식을 이용하여 내부 화소의 움직임을 추정하는 보간 기법 역시 직접 네 개의 꼭지점을 이용하여 표현할 수 있는 장점이 있다.

사각형의 네 꼭지점의 움직임을 알고 있을 때에 내부의 임의의 좌표 x, y 에서는 주변의 움직임을 다항식을 이용하여 보간 할 수 있다. $\boldsymbol{p}=[x,y]^T$ $\boldsymbol{d}(\boldsymbol{p})=[d_x(x,y), d_y(x,y)]^T$ 라 놓으면, 다항식 보간에 의한 움직임은 다음과 같이 표현된다.

$$\boldsymbol{d}(\boldsymbol{p}) = \sum_{\substack{0\le l,m\le N_1 \\ 0\le l+m\le N_T}} \boldsymbol{a}_{l,m}(\boldsymbol{p})\, x^l y^m \qquad (1)$$

영상처리에서 자주 사용하는 affine변환과 bilinear 변환은 식(1)에 대해서 각각 $N_1=N_T=1$ 과 $N_{1=1,}N_T=2$ 로 주어진다. 식(1)에 대한 bilinear 변환식을 전개하면

$$d(\boldsymbol{p}) = a_{00}(\boldsymbol{p}) + a_{01}(\boldsymbol{p})x + a_{10}(\boldsymbol{p})y + a_{11}(\boldsymbol{p})xy$$

로 된다. 이를 마스터 도메인의 네 개 좌표에 반 시계 방향으로 인덱스를 할당하여 식을 전개하면 각 꼭지점의 이동을 나타내는 $d_k(\boldsymbol{p}), k=0,1,2,3$ 에 의하여 식으로 나타난다 [3].

$$d(\boldsymbol{p}) = \frac{1}{4}\sum_{k=0}^{4} \phi_k(\boldsymbol{p})\, d_k(\boldsymbol{p}) \qquad (2)$$

식(1)에 비하여 사각형을 이루는 네 개의 변위를 이용한 식(2)는 동일한 표현이지만 왜곡된 사각형의 모양을 가시화하기 위해서는 식(2)를 사용하여야 한다. 예를 들어, 식(1)에서는 각 변수 값이 주어질지라도 왜곡된 사각형의 모양을 가늠하기 어렵지만, 식(2)는 네 모서리의 움직임에 의한 왜곡 사각형의 모습을 형성하는 것이기 때문에 기하학적인 변형 정도를 쉽게 가시화 할 수 있다. 왜곡 블록 정합 방식은 프레임의 움직임 예측을 위해, 이동 변위 기반의 블록 정합 방식를 이용한다[3].

시간 t_1 과 t_2 에서의 각각의 영상을 $f_{t_1}(\boldsymbol{p})$ 과 $f_{t_2}(\boldsymbol{p})$ 라고 하자. t_1 에 해당하는 프레임을 기준 영상이라 놓으면 t_2 에서의 원 영상과 예측된 영상과의 오차(prediction error)는 식(3)과 같이 정의된다.

$$\mathrm{PE} = \frac{1}{2}\sum_{\boldsymbol{p}\in D}\{ f_{t_1}(\boldsymbol{p}+d(\boldsymbol{p})) - f_{t_2}(\boldsymbol{p})\}^2 \qquad (3)$$

왜곡 블록의 반복적인 움직임 예측이 진행되는 가운데 $t+1$ 번째의 예측영상 $f_{t_1}(\boldsymbol{p}+d^{t+1}(\boldsymbol{p}))$ 를 Taylor 급수를 이용하여 확장할 수 있다.

$$\begin{aligned} f_{t_1}(\boldsymbol{p}+d^{t+1}(\boldsymbol{p})) &= \\ f_{t_1}(\boldsymbol{p}+d^{t}(\boldsymbol{p})) &+ \nabla^T \cdot \sum_{k=1}^{4}\phi_k(\boldsymbol{p})\, d_k(\boldsymbol{p}) \\ &+ \text{Highter Order Terms} \end{aligned} \qquad (4)$$

식(4)에서 ∇^T 는 $[\frac{\partial}{\partial x}, \frac{\partial}{\partial y}]^T$ 를 나타내다. 움직임 예측은 PE 를 최소화하는 움직임 벡터 $d(\boldsymbol{p})$ 를 찾기 위한 것이다. 식(4)를 식(3)에 대입하여 PE=0으로 놓으면 $d_k(\boldsymbol{p})$ 에 대한 선형적인

식으로 정리되어 직접 해를 구할 수 있다.

왜곡 블록 정합 방식의 예측 움직임 벡터는 다음의 행렬의 해로 나타난다.

$$\begin{aligned} \boldsymbol{H}_{xx}\, d_x(\boldsymbol{p}) &= q_x(\boldsymbol{p}) \\ \boldsymbol{H}_{yy}\, d_y(\boldsymbol{p}) &= q_y(\boldsymbol{p}) \end{aligned} \qquad (5)$$

식(5)에서 $\boldsymbol{H}_{xx}$ 와 $\boldsymbol{H}_{yy}$ 는 다음과 같다.

$$e(\boldsymbol{p}) = f_{t_1}(\boldsymbol{p} + d(\boldsymbol{p})) - f_{t_2}(\boldsymbol{p}) \qquad (6)$$

$$\begin{aligned} \boldsymbol{H}_{xx\cdot lk} &= \sum_{p \in D} \left(\frac{\delta f}{\delta x}\right)^2 \phi_l \phi_k \\ \boldsymbol{H}_{yy\cdot lk} &= \sum_{p \in D} \left(\frac{\delta f}{\delta y}\right)^2 \phi_l \phi_k \end{aligned} \qquad (7)$$

$$\begin{aligned} d_x &= [\ d_{1x}\ d_{2x}\ d_{3x}\ d_{4x}]^T \\ d_y &= [\ d_{1y}\ d_{2y}\ d_{3y}\ d_{4y}]^T \end{aligned} \qquad (8)$$

$$\begin{aligned} \boldsymbol{q}_{x,k} &= \sum_{p \in D} e(\boldsymbol{p})\, \phi_k \left(\frac{\delta f_{t_1}}{\delta x}\right), \\ \boldsymbol{q}_{y,k} &= \sum_{p \in D} e(\boldsymbol{p})\, \phi_k \left(\frac{\delta f_{t_1}}{\delta y}\right) \end{aligned} \qquad (9)$$

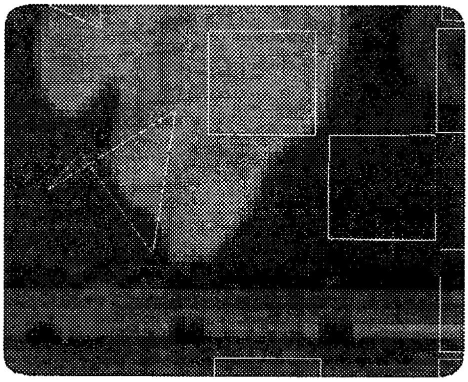
$$\boldsymbol{q}_x = \begin{bmatrix} \boldsymbol{q}_{x,1} \\ \cdots \\ \boldsymbol{q}_{x,4} \end{bmatrix}, \quad \boldsymbol{q}_y = \begin{bmatrix} \boldsymbol{q}_{y,1} \\ \cdots \\ \boldsymbol{q}_{y,4} \end{bmatrix} \qquad (10)$$

$$\begin{aligned} \boldsymbol{H}_{xx} &= \begin{bmatrix} \boldsymbol{H}_{xx\cdot 11} & \cdots & \boldsymbol{H}_{xx\cdot 41} \\ \cdots & \cdots & \cdots \\ \boldsymbol{H}_{xx\cdot 14} & \cdots & \boldsymbol{H}_{xx\cdot 44} \end{bmatrix} \\ \boldsymbol{H}_{yy} &= \begin{bmatrix} \boldsymbol{H}_{yy\cdot 11} & \cdots & \boldsymbol{H}_{yy\cdot 41} \\ \cdots & \cdots & \cdots \\ \boldsymbol{H}_{yy\cdot 14} & \cdots & \boldsymbol{H}_{yy\cdot 44} \end{bmatrix} \end{aligned} \qquad (11)$$

3. 왜곡(歪曲) 블록의 기하학적 오류

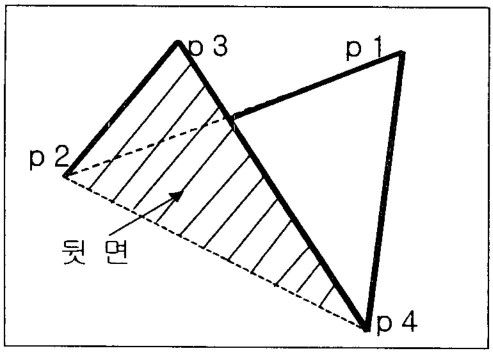
블록 정합 방식에 의한 움직임 예측 기법은 동화상에서 연속적으로 나타나는 물체의 움직임을 예측하여, 중복되는 데이터를 줄이는데 그 목표가 있다. 왜곡(歪曲) 블록 정합 방식 역시 동화상 속에 내포되는 물체의 움직임을 나타내는데 있어서는 동일한 의미를 가지고 있다. 단지 종횡의 움직임만이 아닌 물체의 크기 변화, 회전, 일그러짐 등의 변화에도 올바를 예측을 할 수 있게 그 예측 범위를 포괄적으로 확장하고자 하는 것이다.

왜곡 블록 정합방식은 네 개의 모서리에 대한 움직임을 찾기 때문에, 네 개의 모서리중의 하나의 움직임이 지나치게 되면 모서리 내각이 180도 이상이 되는 경우가 발생한다. 이는 과도왜곡 (over deformation)이라 하며 bilinear mapping의 수치해석을 불안정하게 한다. 또한, 일반적인 두 물체의 움직임에서는 위와 같은 뒤틀림 현상을 배제하기 때문에 수정을 필요로 한다. 움직임 예측 기법이 동화상에 내포된 미세한 움직임을 전제로 하고 있기 때문에 일반적으로는 형태가 크게 변화하지 않고 영상의 움직임을 예측할 수 있다. 하지만, 동화상에 내포된 물체들은 2차원의 공간에 표현됨으로, 이전에 있던 물체가 사라진다거나, 나타나날 수 있다. 이런 물체가 카메라에 투사되어 획득한 영상에 대하여 예측을 진행하기 때문에, <그림 1>과 같이 수학적으로 오류를 나타낼 수 는 과도 왜곡 블록은 기하학적인 오류를 발생하게 된다.



<그림 1> 왜곡 블록 매칭의 오류

4. 왜곡(歪曲) 블록 정합 방식의 과도 왜곡에 대한 검출 및 보정



<그림 2> 기하학적 오류의 예

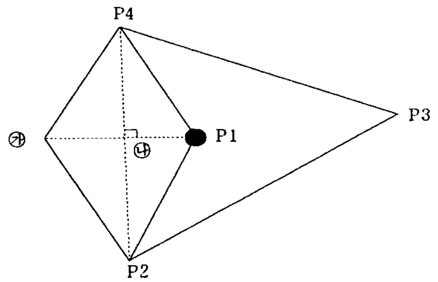
<그림 1>과 같은 과도 왜곡이 발생한 경우를 정리하면 <그림 2>와 같이 영상이 뒤집어지거나 사라지는 영역을 의미한다. 동화상에서 영상이 사라지거나 뒤집어지는 경우 이전 영상에서는 없는 것이 생

긴 것과 같다. 이러한 기학적인 불안정 변화는 자곱비언(jacobian)을 이용하여 확인 할 수 있다[4].

자곱비언은 마스터 도메인에서 x,y 평면으로 기하학적인 변환과정에서 발생하는 면적의 비율을 나타낸다 (식 (12) 참조). 왜곡된 블록의 네 개 이동벡터에 대한 자곱비언을 확인함으로 과도 왜곡을 확인할 수 있다. 자곱비언 함수는 임의의 노드가 뒤집히거나, 한 꼭지점이 두 꼭지점의 일직선상에 놓여서 사라지게 되면 면적의 변화량이 음수가 되거나, 0으로 된다. 왜곡 블록 정합 방식에서 구해진 네 개의 꼭지점에서 자곱비언 함수를 이용한 면적의 변화율을 확인하면 기하학적 오류를 발견할 수 있다.

$$|J| = \left\{\frac{\sum_{k=1}^{4}\phi(s,t)\,x_k}{\delta s}\right\}\left\{\frac{\sum_{k=1}^{4}\phi(s,t)\,y_k}{\delta t}\right\} - \left\{\frac{\sum_{k=1}^{4}\phi(s,t)\,x_k}{\delta t}\right\}\left\{\frac{\sum_{k=1}^{4}\phi(s,t)\,y_k}{\delta s}\right\} \quad (12)$$

예를 들어 <그림 3>에서 네 개의 꼭지점 P1, P2, P3, P4를 가진 사각형을 점검하면, P1에서의 Jacobian은 음수가 되고, ㉯에서는 0으로 나타난다. 곡지점 P1을 내각이 180도가 넘지 않도록 인위적으로 변동시킬 필요가 있다. P3가 움직임 벡터로써의 가치를 잃어 버렸기 때문에 P2와 P4를 연결하는 선과 수직이 되도록 일정거리만큼 떨어진 곳인 ㉮의 위치로 이동하였다. 움직이는 정도는 P1과 ㉯의 간격만큼 이동 시켰으며, 수치해석의 안정성을 위하여 최소거리 이상을 유지하였다. 이를 왜곡 꼭지점의 수직



<그림 3> 왜곡 사각형의 오류 수정

투영이라고 한다.

왜곡 블록의 정합과정에서 Gradient Search 방법을 적용하였으며, 반복적으로 움직임을 찾아가는 과정에서도 계속하여 자곱비언 점검을 필요로 한다. 왜곡 블록 정합 방식의 반복적 작업과정에서 자곱비언 함수를 이용하여 과도 왜곡을 검출해 내고 검출된 과도왜곡을 일정한 규칙에 준하여 수정함으로써 왜곡 블록 정합을 안정적으로 수행 할 수 있다.

5. 결론

왜곡 블록 정합 방식은 기존의 정형화된 블록 정합 방식에 비하여 높은 예측을 보이고 있으며, 본 논문에서는 예측 처리과정에서 발생할 수 있는 기하학적 오류와 그에 대한 정정 방식에 대해 논의 하였다. 자곱비언 함수를 이용하여 마크로 블록(움직임 예측을 위한 물체)의 기하학적 오류를 점검하고 그에 대해 수직 투영시킴으로 과도 왜곡을 정상화 시킬 수 있었다. 하지만, 이는 단지 수학적 기반에 의해 획득되어진 움직임을 임의의 변화시켜 수정하는 방법일 뿐 움직임을 올바로 예측했다고 볼 수는 없음으로 나머지 세 꼭지점을 동시에 수정하는 방법에 대하여 차후 연구가 필요하다. 또는 수직투영에만 의존할 것이 아니라 과도 왜곡이 발생한 블록에 대해서는 4개의 블록으로 분할하여 다시 예측을 수행하는 방법도 고려할 수 있다. 움직임 예측 방법에서의 목표는 예측된 영상이 원 영상과 가능한 동일하도록 만드는 것이다. 또한 본 논문에서는 네 개의 꼭지점을 동시에 이동시키는 방법에 의존함으로 기하학적인 과도 왜곡이 발생하였다. 따라서 각 꼭지점의 움직임 예측을 독립적으로 수행하고 각 모서리의 예측 수행 영역에 제한을 두게 되면 기하학적인 오류는 제거할 수 있을 것으로 판단한다.

[참고문헌]

[1] 고성제, 김종옥 역저, 다양한 영상음성을 자유자재로 부호화하는 MPEG-4의 세계, 영풍 문고, 1999.

[2] Thomas Sikora, The MPEG-4 Video Standard Verification Model, IEEE Trans. on Circuits and Systems for Video Technology, pp.19 - 31, Vol. 7, No. 1, Feb. 1997

[3] O. Lee and Y. Wang, Motion compensated Prediction Using Nodal-Based Deformable Block Matching , Journal of Visual Communication and Image Representation, Vol. 6, No. 1, 1995

[4] O. C. Zienkewicz and R. L. Taylor, The finite element method, vol. 1. Prentice Hall, 4-th ed., 1989.