

HCI 시스템의 손 추적을 위한 수정 블록 정합 알고리즘

The Modified Block Matching Algorithm for a Hand Tracking of an HCI system

김진옥*
Jin-Ok Kim

요약

GUI(Graphic User Interface) 기반의 상호작용은 컴퓨터를 더 사용하기 간단하고 쉽게 만들었다. 그러나 GUI 기반의 상호작용은 자연스럽게 직관적이며 적응적인 사용자의 요구사항을 만족시키기 위해 필요한 상호 작용 기능을 쉽게 지원하지는 못한다. 본 연구에서는 이미지 시퀀스에서 손을 추적하고 가상 현실에서 포인팅 장치로 마우스를 대체하기 위해 각 비디오 프레임에서 손을 인식하는데 유용한 방법인 수정 BMA를 제안했으며 이를 이용해 초당 30 프레임의 HCI 시스템을 구현했다. HCI 시스템을 구현하는데 가장 중요한 기준은 정확한 움직임 벡터 포착과 그의 실시한 처리이다. 수정 BMA는 실시간 처리를 위해 손의 위치, 움직임 방향을 고려한 손 영역을 분할, 손 영역의 색상 분포를 예측하는데 적용했다. 실험 결과는 YCbCr 좌표를 이용한 수정 BMA가 실시간 처리와 인식율을 보장함을 보여 준다. YCbCr 색상 좌표는 각 픽셀 색상의 휘도를 제거한 RGB 색상 좌표보다 더 적은 비트로 코딩 가능하며 주변 상황에 덜 민감하다.

수정 BMA를 이용한 손 추적은 가상현실, 게임과 장애인을 위한 HCI시스템 적용가능하다.

Abstract

A GUI (graphical user interface) has been a dominant platform for HCI (human computer interaction). A GUI -based interaction has made computers simpler and easier to use. The GUI -based interaction, however, does not easily support the range of interaction necessary to meet users' needs that are natural, intuitive, and adaptive. In this paper, the modified BMA (block matching algorithm) is proposed to track a hand in a sequence of an image and to recognize it in each video frame in order to replace a mouse with a pointing device for a virtual reality.

The HCI system with 30 frames per second is realized in this paper.

The modified BMA is proposed to estimate a position of the hand and segmentation with an orientation of motion and a color distribution of the hand region for real -time processing. The experimental result shows that the modified BMA with the YCbCr (luminance Y, component blue, component red) color coordinate guarantees the real -time processing and the recognition rate. The hand tracking by the modified BMA can be applied to a virtual reality or a game or an HCI system for the disable.

Keyword : Hand Tracking, HCI, PUI, Block Matching Algorithm(BMA)

1. 서론

비디오 이미지에서 손의 동작을 인식하는 과정은 단순하면서도 직관적인 HCI(Human Computer Interface)를 제공하는 수단으로 많이 연구되어 왔다[1][2].

최근 포터블컴퓨팅이 각광을 받으면서 키보드가 없는 펜 컴퓨터 분야나 펜컴퓨팅의 핵심 기술

은 손으로 쓴 글자와 동작 인식이 차지하고 있다 [3][4][5].

그러나 이전의 동작 기반의 응용분야 대다수는 동작을 입력하기 위해 마우스, 철틀, 데이터 입력용 장갑과 같은 특별한 하드웨어를 필요로 한다. 현재의 동작 입력 장치가 개선되었다고는 하나 특히 컴퓨터 장치를 이용하는 것이 두려운 초보자에게는 특별한 하드웨어를 이용하여 동작을 생성하는 것이 자연스럽게 지 못한 일이다. 본 연구에서는 손 영역을 추출하여 동작을 인식하고 인식

* 정회원 : 세종대학교 컴퓨터공학과 초빙교수
jokim@sejong.ac.kr(제1저자)

결과를 이용하여 윈도우 매니저를 제어하는 방법을 제안한다. 손 이미지를 추출하는데에는 피부색상 모델의 색상이미지 처리와 색상의 레이블링을 이용한다. 손의 궤적은 입력 이미지 시퀀스의 각 프레임으로부터 수정된 BMA(Block Matching Algorithm)을 이용하여 얻는다[6][7][8]. 궤적이 미리 설정된 윈도우 관리규칙에 따라 동작을 인식하기 위해 분석될 때 동작 인식 기술은 윈도우 매니저가 디스플레이를 제어하는데 이용한다[9][10][11].

2. 손 영역 분할

손 영역 분할 알고리즘은 빠른 저수준 영역 분할과정에서 색상정보 사용을 포함한다[12][13]. 이의 목적은 입력 이미지의 픽셀을 피부 색상과 그 외 색상을 구분하기 위해서이다[14][15][16][17]. 피부 색상 참조 지도는 YcbCr 색상 공간에서 이용된다.

피부 색상 영역은 YCbCr 색상 공간에서 색차 신호 집합(i.e., Cb와 Cr)의 값이 지속적으로 나타나는지 여부를 이용하여 식별한다. [Cb] 범위와 [Cr] 범위는 결과적으로 피부 색상 참조 지도에 명시된 피부 색상에 해당하는 Cb와 Cr 값의 개별 범위를 의미한다.

비디오 이미지에서 감지한 다른 피부 색상은 이미지 영역의 색차와 다르지 않다. Cb와 Cr값에서 파생한 이런 지도는 피부 색상 변화와 상관없이 유효하다. 인종간의 피부 색상의 차이는 주로 피부가 밝고 어두운 정도의 차이이다. 이러한 특성은 Y에 의해 결정되는 색상의 밝기 차이에 의해 구분되기 때문에 Cr과 Cb와는 관계가 없다.

피부 색상 참조 지도를 이용하여 색상의 분할은 식(1)과 같이 이루어진다

분할은 입력 이미지의 색차 성분만을 필요로 한다.

$$Out = \begin{cases} 1, & \text{if } [C_{(x,y)}] \in Range(C_r) \cup [C_{(x,y)}] \in Range(C_b) \\ 0, & \text{otherwise} \end{cases}$$

(1)

출력 픽셀은 피부 색상으로 분류되고 픽셀의 Cb 또는 Cr 값이 고려 범위에 있으면 1로 세팅되며 그렇지 않으면 픽셀은 피부 색상이 아니라고 구분되어 0값이 세팅된다.

가장자리의 흐릿함을 제거하기 위해 Median 필터를 사용하여 주변에서 밝기 중간값으로 현재 이미지 포인트를 대체한다. 주변의 밝기 중간값은 개별 스파이크 잡음 영향을 주지 않기 때문에 Median 필터는 임펄스 잡음을 어느 정도 제거할 수 있다. 또한 Median 필터링은 가장자리를 흐릿하게 만들지 않기 때문에 반복하여 적용가능 하다.

3. 움직임 예측

3.1 블록 정합 알고리즘(BMA)

블록 정합 알고리즘은 움직임 예측과 추적에 많이 이용되는 상관성 기반의 접근 방법이다. 손 동작의 방향과 궤적은 분포도와 추출된 움직임 벡터의 방향에 따라 유도한다. 그래서 움직임 벡터의 분포도는 통계적으로 처리되어 손 동작 인식에 적용된다.

이미지는 각 블록내의 전이 움직임은 동일하다는 전제하에 중첩하지 않은 집합으로 나뉘어 동일하게 사각형 블록으로 간격 처리된다.

이 단순한 모델은 전이 움직임만을 고려하지만 큰 대상의 회전과 확대,축소와 같은 다른 움직임은 충분히 작은 블록들의 구분적인 전이에 의해 근사치화 할 수 있다.

이 블록들의 움직임 벡터는 이전 프레임에서 최적의 정합 대상을 찾음으로써 예측한다. 각 블록의 움직임은 한 개의 움직임 벡터로만 설명되기 때문에 움직임 벡터의 측면 정보는 줄고 사각형 모양 정보는 미리 인코더와 디코더 양쪽이 알고 있어야 한다.

블록의 크기는 적절하게 선택되어야 한다. 일반적으로 크기가 작을수록 근사값이 더 정확하게 계산된다. 그러나 작은 블록 크기는 더 많은 움직임

벡터를 만든다. 조정결과 16×16 를 적절한 크기로 결정한다. 이 내용은 H.261, H.263, MPEG-1, MPEG-2 등의 국제 비디오 코딩 표준에 기술되어 있다.

3.2 정합 기준

블록의 정합은 최대 상관성, 최소 MSE(mean square error), 최소 MAD(mean absolute difference), 최대 MPC(matching per count)등을 포함한 다양한 기준에 따라 정량화한다.

최소 MSE 기준에서 MSE는 식(2)와 같이 정의한다.

$$MSE(d_1, d_2) = \frac{1}{N_1 N_2} \sum_{n_1, n_2 = \beta} [F_k(n_1, n_2) - F_{k+1}(n_1 + d_1, n_2 + d_2)]^2 \quad (2)$$

여기서 $MSE(d_1, d_2) = \frac{1}{N_1 N_2} \sum_{n_1, n_2 = \beta} |F_k(n_1, n_2) - F_{k+1}(n_1 + d_1, n_2 + d_2)|^2$ 는 예상 움직임 벡터의 집합 (d_1, d_2) 에 $N_1 \times N_2$ 대한 블록이다. 움직임 벡터의 예측은 식 (3)과 같이 MSE를 최소화 (d_1, d_2) 하는 값을 취한다.

$$[d_1, d_2]^T = \operatorname{argmin} MSE(d_1, d_2) \quad (3)$$

MSE 기준의 최소화는 블록의 모든 픽셀에 대한 광류 제약으로 볼 수 있다. 치환된 프레임 차이가 표현은 식 (4)와 같다.

$$df_{d_k, d_{k+1}}(n_1, n_2) = [F_k(n_1, n_2) - F_{k+1}(n_1 + d_1, n_2 + d_2)] \quad (4)$$

그러나 최소 MSE 기준은 하드웨어에서 제품 연산을 구현하기가 어렵기 때문에 일반적으로 하드웨어 구축에서는 사용되지 않는다. 대신에 식 (5)의 최대 MAD 기준이 실제적으로 많이 사용된다.

$$MSD(d_1, d_2) =$$

$$\frac{1}{N_1 N_2} \sum_{n_1, n_2 = \beta} |F_k(n_1, n_2) - F_{k+1}(n_1 + d_1, n_2 + d_2)| \quad (5)$$

3.3 수정 블록 정합 알고리즘

BMA가 각 픽셀의 이동량을 가정하는 것은 블록에서 움직임을 가정하는 것과 동일하며 이동량은 검색 윈도우 크기 한계를 넘어서면 찾을 수 없기 때문에 이러한 가정은 블록킹 효과와 이동량 측정 오류의 원인이 된다. 그러므로 블록킹 효과와 움직임 벡터의 예측 오류를 보상하기 위해 검색 윈도우의 크기를 더 적절하게 정해야 한다.

100개의 랜덤 변수가 검색 윈도우에서 만들어지고 상관 윈도우의 중앙 점이 각 포인트에 위치하면 먼저 MAD가 가진 최소값의 포인트를 찾을 수 있다. 이 방법에 따라 검색 윈도우가 16×16 크기로 만들어지면 50개의 랜덤 변수는 최적의 움직임 벡터에 고정된다. 이렇게 하면 초기값에 의해 야기된 움직임 변수 위치와 같은 문제점들은 이 알고리즘에 의해 해결되고 적절한 분포 형태를 찾아 낼 수 있다.

BMA 알고리즘은 모든 픽셀의 이동이 블록에서 같이 이루어진다는 가정하게 적용되기 때문에 특정 부분의 움직임을 찾기가 어렵다. 그러나 손 궤적 추적에서 필요로 하는 포인트의 지정과 포인트의 제동은 손가락의 끝점들에서만 이루어진다. 손가락의 끝과 손등간의 차이는 손가락 끝을 찾음으로써 분류 가능하다. 어떤 특정 점이 지정되었을 때 손가락 끝의 이동은 관계적으로 손등의 움직임보다 크다. 움직임 벡터의 크기가 히스토그램에서 먼저 표현된 후에 분포와 임계값에 적절한 값이 고정된다. 결국 이러한 값들이 계산되면 손 끝의 움직임 벡터가 포착된다. 즉 손가락 끝의 움직임이 포인트로, 손은 제동자로 인식된다. 인식 방법은 움직임 벡터 이동량의 히스토그램으로 표현된다.

4. 실험

4.1 손 영역 분할

YCbCr 색상 좌표는 비디오 신호에 대한 국제 표준이고 BT878 칩은 YCbCr(4:2:2)의 색상 포맷을 생성한다. 그러므로 RGB를 YcbCr로 변환하는 계산이 필요없기 때문에 처리 시간은 줄어든다. 히스토그램에서 Cb와 Cr의 값은 손 영역의 전체 이미지를 분할한다. 세사람에 대해 각각 10개의 프린트를 했기 때문에 실험에서는 Cb와 Cr 값은 30개의 프린트를 이용하여 식 (6)과 같이 추출한다.

$$88 \leq C_b \leq 120, \text{ and } 130 \leq C_r \leq 160 \quad (6)$$

Cb와 Cr값을 갖는 BMA 분할의 경우 잡음 때문에 잘못된 움직임 벡터를 찾을 수 있기 때문에 수정 BMA는 움직임 벡터를 얻는데 이용한다. 입력

이미지 크기는 320×240이고 초기 블록 크기는 16×16 픽셀이다. 검색 윈도우 크기는 48×48 픽셀이다. 초기 검색 윈도우에서 100개의 변수가 생성되고 16×16 크기의 검색 윈도우가 윈도우 중앙에 만들어진다. 최종 움직임 벡터는 추가 50개 변수를 이용하여 찾아낸다.

실험결과는 수정BMA가 계산시간과 픽셀 검색에서 우수한 특성을 지녔음을 보여준다.

이 과정은 본 연구에서 손가락 움직임 마우스를 실시간 구현하는데 충분하다.

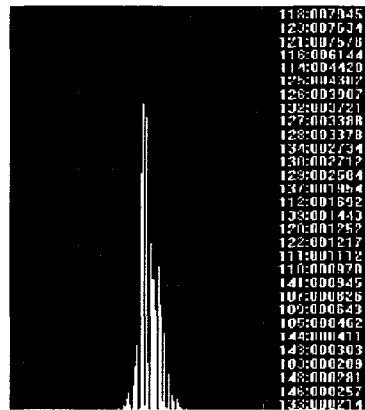
그림 1은 손 영역과 그에 해당하는 YCbCr의 히스토그램이고 그림 2는 손 영역과 그의 윤곽선이다.

4.2 동작 인식

HCI 시스템에서 실시간 처리와 온라인 처리가 이루어지지 않으면 손가락 동작으로 이루어지는 일은 무의미하다. 그러므로 현재 사용되는 마우스보



(a) 손 영역

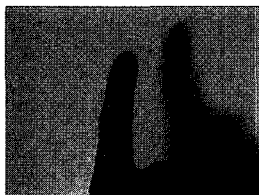


(b) 손 영역의 Cb

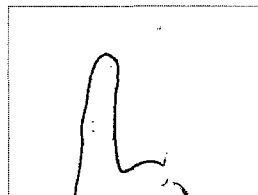
(그림 1) 손영역과 YCbCr의 히스토그램



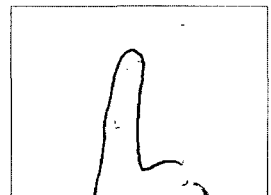
(a) 원 이미지영상 F(t-1)



(b) 원 이미지 영상 F(t)

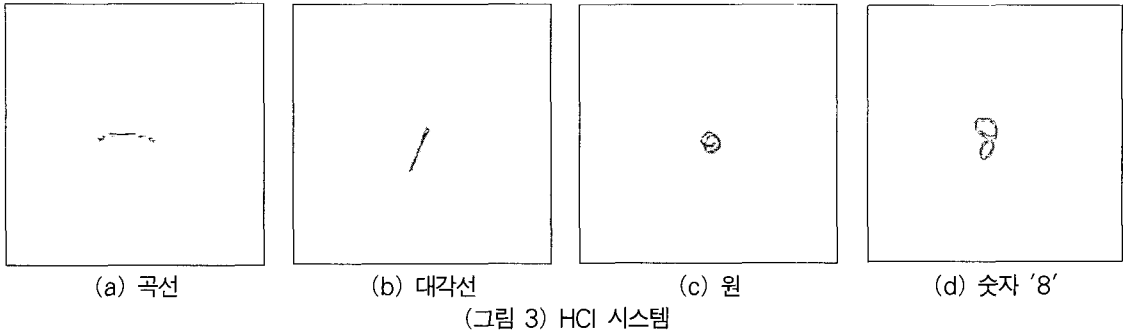


(c) F(t-1)의 윤곽선



(d) F(t)의 윤곽선

(그림 2) 손 영역과 윤곽선



다 계산 시간이 더 걸린다면 알고리즘의 성능과는 무관하게 해당 응용 방법은 효율성이 떨어진다.

30프레임을 실행하는 시스템을 실험을 위해 구축하여 움직임 벡터의 결과를 HCI 시스템에 적용했다. 실험 결과는 손 영역 분할과 YCbCr 색상 좌표를 이용한 움직임 벡터 포착 그리고 수정 BMA가 실시간, 온라인 처리를 위해 구현될 수 있음을 보여 준다.

이 방법의 HCI 시스템은 엄지손가락이 펴지지 않거나 움직이지 않을 때 포인팅 장치로, 모든 손가락이 펴지지 않을 때는 제동 장치로 이용 가능하다.

또한 근소한 움직임에 동작하는 것을 피하기 위해 20 이하의 움직임 벡터에서 작동하도록 설정하며 포인팅과 클릭 사이의 움직임 벡터의 방향 차이와 그 크기를 인식한다.

그림 3은 수정 BMA를 적용한 HCI 시스템이 어떻게 작동하는지 보여 준다. 포인팅하고 움직일 때 모니터는 그 결과를 추적한다. 클릭했을 때 시스템은 클릭 동작을 실행한다.

새로운 클릭이 모든 이전의 흔적을 지우고 초기 화면으로 되돌아가면 포인터는 디스플레이 장치의 중앙에 위치한다.

5. 결론

GUI(Graphic User Interface) 기반의 상호작용은 컴퓨터를 더 사용하기 간단하고 쉽게 만들었다. 그러나 GUI 기반의 상호작용은 자연스럽고 직관적이며 적응적인 사용자의 요구사항을 만족시키

기 위해 필요한 상호 작용 기능을 쉽게 지원하지는 못한다. 본 연구에서는 이미지 시퀀스에서 손을 추적하고 가상 현실에서 포인팅 장치로 마우스를 대체하기 위해 각 비디오 프레임에서 손을 인식하는데 유용한 방법인 수정 BMA를 제안했으며 이를 이용해

초당 30 프레임의 HCI 시스템을 구현했다. HCI 시스템을 구현하는데 가장 중요한 기준은 정확한 움직임 벡터 포착과 그의 실시한 처리이다. 수정 BMA는 실시간 처리를 위해 손의 위치, 움직임 방향을 고려한 손 영역을 분할, 손 영역의 색상 분포를 예측하는데 적용했다. 실험 결과는 YCbCr 좌표를 이용한 수정 BMA가 실시간 처리와 인식율을 보장함을 보여 준다. YCbCr 색상 좌표는 각 픽셀 색상의 휘도를 제거한 RGB 색상 좌표보다 더 적은 비트로 코딩 가능하며 주변 상황에 덜 민감하다.

수정 BMA를 이용한 손 추적은 가상현실, 게임과 장애인을 위한 HCI시스템 적용에 가능하다.

참고문헌

- [1] Mitchell, J. L., Pennebaker, W. B., Fogg, C. E., LeGall, D.J.: MPEG Video Compression Standard. Champion and Hall, 1996.
- [2] Haskell, B. G., Puri, A., Netravali, A. N.: Digital Video: An Introduction to MPEG -2. Champion and Hall, 1997.
- [3] Shi, Y. Q., Sun, H.: Image and Video Com-

- pression for Multimedia Engineering. CRC Press, 2000.
- [4] Russ, J.C.: The Image Processing Handbook. 3 edn. CRC Press, 1999.
- [5] Bae, C. S., Chun, B. T., Min, B. W.: Window manager control by extraction hand region and analysing its trajectory. Journal of KISS 24, 1997.
- [6] Ahmad, T., Taylor, C. J., Lantitis, A., Cootes, T. F.: Tracking and recognition hand gestures using statistical shape model. Image and Vision Computation 15, 1997.
- [7] Crane, R.: A Simplified Approach to Image Processing. Prentice Hall, 1997.
- [8] Parker, J. R.: Algorithm for Image Processing and Computer Vision. Wiley Computer Publishing, 1996.
- [9] Tekalp, A.M.: Digital Video Processing. Prentice Hall, 1995.
- [10] Bovik, A.: Handbook of Image and Video Processing. Academic Press, 2000.
- [11] Pratt, W. K.: Digital Image Processing. 3 edn. John Wiley, 2001.
- [12] Sonka, M., Hlavac, V., Boyle, R.: Image Processing, Analysis, and Machine Vision. PWS Publishing, 1999.
- [13] Gonzalez, R. C., Woods, R. E.: Digital Image Processing. 2 edn. Addison - Wesley, 2002
- [14] Forsyth, D. A., Ponce, J.: Computer Vision: A Modern Approach. Prentice Hall, 2002.
- [15] Duda, R. O., Hart, P. E., Stork, D. G.: Pattern Classification. 2 edn. Wiley - Interscience, 2000.
- [16] Abe, S.: Pattern Classification. Springer Verlag, 2001.
- [17] Javidi, B.: Image Recognition and Classification. Marcel Dekker, 2002.

● 저 자 소 개 ●



김진욱

1989년 성균관대학교 졸업 (학사)

1998년 성균관대학교 정보통신공학과 졸업(석사)

2002년 성균관대학교 전기전자 및 컴퓨터 공학과 졸업(박사)

2003년~현재 : 세종대학교 컴퓨터공학과 초빙교수

1992년~1994년 (주)현대전자산업 정보통신사업본부

1994년~1999년 (주)현대정보기술 인터넷사업본부 과장

1999년~2000년 (주)은세통신 온라인사업 팀장

2000년~2001년 (주)유로코넷 기술담당 이사

관심분야 : Multimedia Processing, Image Processing, Biometrics, Recognition

E-mail : jokim@sejong.ac.kr