

포인터의 패턴인식을 통한 프레젠테이션 제어

조동현*장희정 김계영 최형일
숭실대학교 컴퓨터학과

{tluke, heeid}@vision.ssu.ac.kr {gykim, hic}@computing.ssu.ac.kr

Presentation Control Using Pattern of Pointer Recognition

Dong-Hyun Cho Hee-Jung Jang Gye-Yong Kim Hyung-II Choi

Department of Computing, SoongSil University

요 약

각 단체나 직장등에서 프레젠테이션은 필수적인 요소이다. 그리고 근래에는 프로젝터의 보급률이 늘어나면서 과거의 OHP필름을 사용하는 방법보다는 직접 컴퓨터를 프로젝터에 연결하여 자료를 영사하면서 프리젠테이션을 진행하는 사례가 늘어나고 있다. 이러한 환경에서 사용자는 컴퓨터의 제어에 있어서 거리라는 제약을 받아야만 했다. 프리젠테이션을 진행하는 동안 사용자는 컴퓨터를 제어할 수 있는 위치에서 발표를 해야 하지만 레이저포인터라는 일반적인 도구를 사용하여 컴퓨터에 명령을 내릴 수 있다면 더 효율적으로 거리에 구애받지 않고 자료의 이동이나 다른 작업을 수행할 수 있다. 또한 컴퓨터가 학습하는 패턴을 늘려감으로서 더욱 다양한 명령을 지시할 수 있게 함으로써 프레젠테이션환경의 개선 할 수 있다.

1. 서 론

최근 빔 프로젝터의 보급이 늘어나면서 발표자는 직접 프레젠테이션 자료를 컴퓨터로 제어하면서 발표하는 경우가 늘어나고 있다. 이 경우 발표자는 프레젠테이션을 진행하는 동안 컴퓨터의 근거리에 머물러야 하는 제약이 있으며 원거리에 있더라도 발표 자료의 페이지 이동 등의 명령시에 컴퓨터를 직접 제어할 수 있는 근거리로의 이동이 필요하게 된다. 본 논문에서는 발표자가 컴퓨터의 근거리로의 이동 없이도 발표환경 어디에서나 발표 자료를 제어할 수 있는 시스템의 구현을 목적으로 한다. 본 논문에서 제안하는 방법은 시중에 많이 유통되고 프레젠테이션 시에 대부분의 발표자가 사용하는 레이저 포인터와 PC용 카메라를 사용하여 컴퓨터를 제어한다. 레이저포인터로 특정 영역을 지정하여 컴퓨터를 제어하는 경우에는 그 동작이 극히 제한적이고 오작동의 확률도 높아진다. 이것을 방지하기 위해 포인터로 특정 패턴을 그리게 되면 그 패턴을 인식하여 사용자가 원하는 명령의 패턴인지 아니면 단순한 포인팅인지를 판별하여 사용자가 원하는 명령을 컴퓨터에 전달하게 된다. 이를 수행하기 위하여 본 논문에서는 히든 마코프 모델(Hidden Markov Model)을 사용하여 인식률을 높이고자한다.

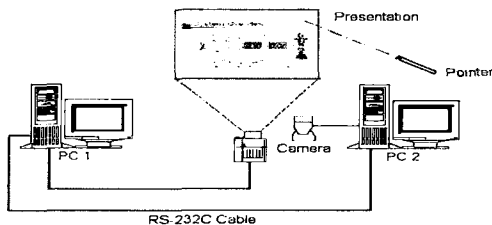


그림1 전체적인 시스템 구성도

그림1은 본 논문에서 제안하는 전체적인 시스템 구성도

이다. 그림1과 같이 PC1은 PC2로부터 명령을 받아들이 실제 프레젠테이션을 제어하는 기능을 하고 PC2는 카메라로부터 받아들인 영상에서 포인터를 추적하여 그 패턴을 인식하는 역할을 한다. 이때 카메라는 프레젠테이션 영역을 촬영하게 되며 사용자는 프레젠테이션 화면에 패턴을 그림으로써 제어를 위한 명령어를 전달한다. 본 논문은 총 3장으로 구성되어 있다. 2장에서는 레이저 포인터를 배경영역과 분리해내기 위해 사용한 YIQ 컬러모델과 배경으로부터 포인터의 좌표를 추출하는 과정을 설명하고, 3장에서는 실제적으로 패턴을 인식하기 위한 과정으로 방향코드의 추출방법과 추출된 방향코드의 집합을 이용하여 HMM에 적용시키는 방법을 대해서 기술한다. 4장에서는 실험결과, 5장은 결론 및 향후 연구 과제를 기술한다.

2. 포인터와 배경의 분리

레이저포인터는 프레젠테이션 되고 있는 화면보다는 확연히 다른 색상 값을 가진다. 특히 휘도가 매우 높은 붉은 색을 띄기 때문에 프레젠테이션 내에 다른 물체가 들어왔을 때에도 그 색상 값의 변화는 레이저포인터의 색상 값의 차이를 넘지 못한다. 따라서 포인터의 분리는 배경의 변화에 임계치를 설정하여 레이저포인터의 입력 유무를 분리해 낼 수 있다.

2.1 YIQ에서 배경의 모델링

인간은 컬러의 구별에 밝기와 색상, 채도의 세 가지 특징을 이용한다. 본 논문에서는 이러한 특징을 잘 나타내는 HSI나 YIQ 컬러공간 중 밝기의 변화에 있어서 색상이나 채도가 변하지 않는 YIQ 컬러모델을 사용한다[1]. 일반적으로 프로젝터의 조도는 500 ~ 5000 ANSI Lumen 인데 반해 레이저포인터의 조도는 10000 ANSI Lumen 이상의 휘도가 매우 높은 노란색이다. 그러므로 임으로 밝기 값과 색상 값의 적절한 임계치를 사용하면 손쉽게

포인터영역을 분리해 낼 수 있다. 이러한 조건으로 배경 영역에 대한 일정시간 동안의 프레임의 평균하여 배경영역을 모델링 한다. 본 논문에서는 YIQ 컬러모델을 사용하므로 평균되어진 배경의 값들은 식(1)의 형태로 저장된다. 이때 표준편차 식(2)도 구해진다. 배경영역의 평균과 표준편차를 구하는 식은 식(3), (4)에 보이고 있다. 저장된 배경과 입력된 영상을 비교하는 식은 다음 식(3)과 같다. 여기서 I는 입력영상을 의미한다. 이 식에서 구해진 D가 임계치를 초과하면 포인터로 간주하고 위치를 추적한다.

$$(\mu_Y, \mu_I, \mu_Q) \tag{1}$$

$$(\sigma_Y, \sigma_I, \sigma_Q) \tag{2}$$

$$\mu_a = \frac{\sum_{x=0}^M \sum_{y=0}^N f_a(x,y)}{M \times N} \quad \text{where } a = Y, I, Q \tag{3}$$

$$\sigma_a = \sqrt{\frac{\sum_{x=0}^M \sum_{y=0}^N (f_a(x,y) - \mu_a)^2}{M \times N}} \quad \text{where } a = Y, I, Q \tag{4}$$

$$D_i(x) = \sum_{i=1,2,3} [(I_i(x) - \mu_i(x)) / \sigma_i]^2 \tag{5}$$

2.2 배경과 포인터 영역

배경영역이 모델링 되면 배경영상과 포인터영상을 분리해 내야 한다. 본 논문에서는 배경영상의 일반적인 변화는 포인터 입력을 넘지 못하기 때문에 임의의 시간(t) 동안의 프레임의 평균값과 새로 입력받은 영상을 비교하여 포인터의 영역을 추출하는 방법을 사용한다.

3. 레이저 포인터의 패턴인식

3.1 포인터의 패턴인식

레이저포인터의 패턴인식의 과정은 다음과 같다. 첫째 입력된 영상에서 포인터의 영역을 분리한다. 탐색영역에 레이저포인터가 들어오는 그 순간부터 좌표를 입력받고 각 좌표들은 시간의 순서대로 연결된다. 그리고 이 좌표의 연속은 특정한 패턴을 갖는다. 명령의 인식을 위해서는 사용자가 인위적이 아니고서는 만들어내기 어려운 패턴(예 : 원, 삼각형, 사각형 등)을 지정하여 그 패턴을 인식하도록 하는 방법을 사용한다. 패턴인식을 위한 수행 과정은 그림 2와 같다[2].

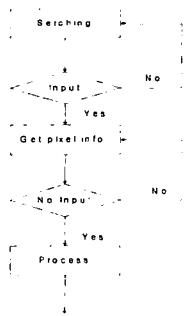


그림2 패턴인식을 위한 시스템 흐름도

둘째 분리된 영역의 좌표 값을 추출한다. 포인터는 다른 이동물체보다 그 영역이 매우 작고 모양의 변화가 거의 없어 간단한 알고리즘으로 좌표값을 추출할 수 있다. 그러나 입력된 점을 확대해 보면 주위로 빛이 분산되는 등 10 Pixel정도의 영역에 영향을 미치게 된다. 또한 발표자가 측면에서 지시를 할 경우 등에는 영향을 미치는 영역이 더 넓어지는 현상을 보인다. 그래서 대표 값을 정해야 한다. 본 논문에서는 임계치 이상의 값들 중 가장 먼저 입력되는 값을 그 포인터의 좌표를 입력하는 값으로 본다. 셋째 패턴을 구성한다. 그러나 포인터만을 가지고 사용자가 원하는 곳에 정확히 원하는 패턴을 그리기는 쉽지 않다. 일반적으로 입력된 영상은 탐색영역 밖으로부터 안으로 들어온 후 패턴을 그리게 된다. 이 경우가 그려진 패턴은 다음절에서 설명할 방향코드를 얻기에 부적절한 모양의 패턴이 그려지게 된다. 이 경우에 적절한 패턴만을 추출하기 위한 알고리즘이 필요하다. 그리고 본 논문에서는 폐곡선의 도형을 패턴으로 사용하기 때문에 필요 없는 부분을 배제해야 한다. 이 문제를 해결하기 위해 입력된 패턴의 끝점을 기준으로 교차하는 점이 있는지를 판별하여 교차점이 있는 경우에는 교차된 내부의 패턴만을 사용하고 교차점이 없는 경우에는 가장 가까운 점을 연결하여 폐곡선으로 만든다.

3.2 방향코드

패턴 인식을 위하여 본 논문에서는 문자인식에서 많이 사용되고 있는 방향코드를 사용한다[3,4]. 방향코드란 특정 패턴에 대하여 일정 방향에 대하여 번호를 매긴 것이다. 즉 패턴 O가 입력되었을 때 패턴 O를 임의의 수 T로 나누고 각각의 t0, t1, t2, ..., tn을 연결하여 각 구간의 선분의 각도를 측정하여 소속하는 코드화 시키는 방법을 말한다. 일반적으로는 360°를 45°씩 분할하여 0 ~ 7까지의 코드를 취하는 8방향코드가 많이 사용되고 있다. 그런데 사람이 펜으로 그린 패턴은 부드럽지만 포인터 사용하여 그린 패턴은 매우 구불구불하며 불규칙적이다. 본 논문에서는 이런 경우에 대해 인식률을 높이기 위해서 8방향코드를 22.5°회전시킨 코드를 사용한다.

3.3 HMM의 적용

본 논문에서는 입력받은 패턴을 20개의 구간으로 나눈 후 방향코드를 추출한다. 이 방향코드들은 Baum-Welch 알고리즘에 의해 학습된 방향코드와 비교되어 어느 클래스에 속할 것인지를 평가받게 된다.

4. 실험 및 결과

본 논문에서는 5명의 실험자가 그린 20개의 패턴을 입력받아 실험을 하였다. 각각의 실험자들에게 지정된 패턴인 원, 사각형 그리고 의미 없는 패턴을 그리게 하여 정확도를 측정해 보았다. 그리고 임의의 물체(예 : 손, 팔, 머리 등)를 탐색영역에 삽입했을 때 포인터영상과의 구분여부를 관찰 해 보았다. 마지막으로 기존의 방향코드를 회전시켜서 같은 방법으로 실험해 보았다. 그림 3은 실제 실험 장면을 보여준다.

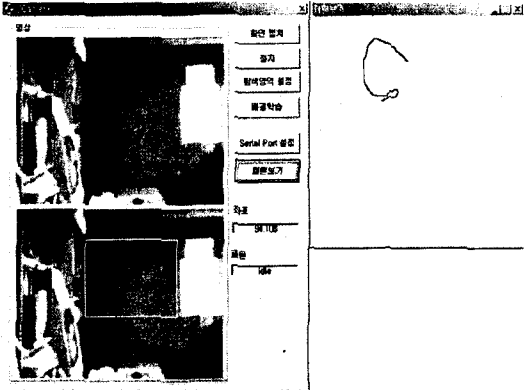


그림3 실제 실험장면

실험결과 원의 경우에는 패턴이 변해 가는 추이가 거의 비슷하기 때문에 인식률이 매우 높게 나타났다. 또한 아무 의미 없는 패턴의 경우에도 인식률이 매우 높게 나타났다. 그러나 사각형의 경우는 사람마다의 특성에 따라서 방향코드의 편차가 심하기 때문에 인식률이 저조하게 나타났다. 그리고 패턴을 너무 빨리 그리는 경우에는 시스템의 속도상 충분한 방향코드를 얻을 수 없기 때문에 인식이 제대로 이루어지지 않았다. 그림 4는 실험 결과를 보여준다.

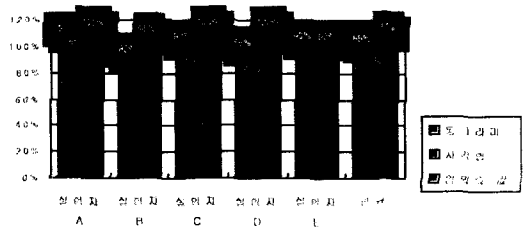


그림 4 실험결과

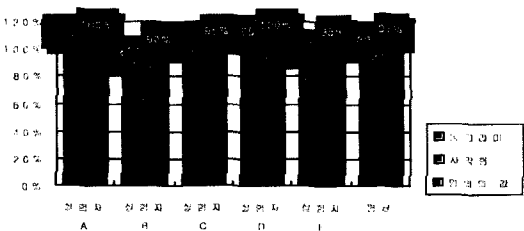


그림5 22.5. 회전한 방향코드를 이용한 결과

그림5에서는 기존의 방향코드를 22.5°회전한 후 그림 4와 동일한 방법으로 실험을 한 결과이다. 실험결과 원은 그런 패턴에 비해 사각형을 그런 패턴의 인식률이 많이 향상되었음을 알 수 있다. 이는 직선을 그릴 때 학습패턴이나 입력패턴 모두 원래의 방향코드보다 단일한 패턴이 입력되었음을 보여준다.

5. 결론 및 향후 연구과제

본 논문에서는 원거리에서 사용자가 컴퓨터에 원하는 명령을 입력함으로써 직접 컴퓨터에 키보드나 마우스로 입력하는 번거로움 없이 프레젠테이션을 할 수 있는 시스템을 제안했다. 그런데 본 논문에서 제안한 시스템은 카메라로부터 입력받은 영상의 패턴을 이용하여 명령을 인식해야 하기 때문에 컴퓨터의 처리속도에 매우 민감하다. 이것은 HMM의 계산과정과 컴퓨터간의 통신으로 발생하는 지연시간에 원인이 있다. 그래서 영상의 처리과정을 개선하고 컴퓨터간의 통신에 TCP/IP를 사용하면 속도의 향상을 기대할 수 있을 것으로 보인다. 향후 연구과제로는 학습패턴의 확장을 통한 다양한 명령 제어와 프레젠테이션 화면을 칠판처럼 사용할 수 있는 연구가 필요하다.

Acknowledgement

본 논문은 첨단기술정보연구센터(AiTrc)를 통하여 과학재단의 일부를 지원 받았음

참고 문헌

- [1]Rafael C. Gonzalez, Richard E. Woods, Digital Image Processing Addison-Wesley, 1992.
- [2]David Harwood, Larry Davis, Alex Pentland, Real-time Motion Capture, IEEE, 1998.
- [3] L. R. Rabiner, A tutorial on hidden markov models and selected applications in speech recognition, *Proceedings of the IEEE*, Vol. 77, No.2 pp 257-286,1989.
- [4]박 희 선, 이 성 환, “은닉 마르코프 모델을 이용한 필기체 한글의 오프라인 인식” 한국 정보과학회 논문지, 1994.