

Nonstationary HMM 을 이용한 다중 카메라 기반 장소 인식

Multi-Camera-based Place Recognition using Nonstationary HMM

민경민, Kyungmin Min*, 이성훈, Seong Hun Lee*, 김동호, Dongho Kim*, 김진형, Jin Hyung Kim*

*한국과학기술원 전자전산학과 전산학전공

요약 사용자가 현재 위치해 있는 장소를 알아내는 것은 상황인식 분야에서 활발히 연구되고 있는 분야중 하나로, 이를 위해서 사용자의 몸에 다양한 센서를 장착하고 그 센서로부터 추출되는 데이터를 분석하여 사용자의 위치를 인식하는 연구가 많이 이루어져왔다. 본 논문에서는, 사용자의 몸에 장착된 카메라로부터 얻어진 영상을 이용하여 사용자의 현재 장소를 인식하는 장소 인식 시스템을 제안한다. 기존의 방법론들에 비해서 높은 성능을 보이기 위해서 본 논문에서는 두 가지 방법을 제안하였다. 먼저 한 방향만의 영상으로는 인식이 어려운 장소에서도 좋은 인식 성능을 보일 수 있도록 하기 위해, 여러 대의 카메라를 동시에 사용하여 여러 방향의 영상을 얻어내는 방법을 제안하였다. 또한 이전 시간의 장소 인식 결과로부터 현재 시간의 장소를 추론하는 데에 있어서, 각 장소들에 대해 알고 있는 사전지식을 보다 많이 적용할 수 있는 인식 모델을 제안하였다. 실제 대학 실내 환경에서의 실험을 통하여, 제안한 방법을 이용한 장소 인식기법이 좋은 성능을 보임을 확인할 수 있었다.

핵심어: 상황 인식, 장소 인식, 영상 이해, 웨어러블 컴퓨팅, HMM

1. 서론

유비쿼터스 컴퓨팅에 대한 연구가 활발해지면서, 사용자에게 장착된 센서로부터 사용자의 현재 상황을 인식하고, 그에 따라 적합한 서비스를 제공해주는 상황 인식 서비스에 대한 관심이 증대되고 있다.

사용자의 상황에 적합한 서비스를 제공하기 위해서는, 무엇보다도 사용자의 상황을 인식하는 것이 필수적이다. 인식해야 할 상황은 제공하려는 서비스에 따라 사용자의 위치, 행동, 신체 상태 등으로 다양하게 정의될 수 있다. 그 중에서도 사용자가 현재 위치한 장소를 파악하는 것을 장소 인식이라 한다. 장소 인식을 통해 복도, 강의실, 회의실 등과 같이 실시간으로 인식된 사용자의 위치 정보는 그 자체로도 맵인 보조, 병원 내 환자 위치 추적 등의 다양한 서비스를 제공하는 데 사용될 수 있고, 또한 사용자의 현재 행동을 추론하기 위한 기본 정보로도 사용이 가능하다.

현재까지 장소 인식을 위하여 사용되는 주요한 접근 방법으로는 GPS와 같은 전역적 신호시스템을 이용하여 인식하는 방법, 인식하려는 장소에 RFID 태그를 부착하여 인식하는 방법, 그리고 사용자에게 센서를 장착하여 인식하는 방법이 있다. 그러나 GPS의 경우 정밀도가 낮고, 실내에서는 사용이 어렵다는 단점이 있고, RFID의 경우는 설치 비용이 크다는 단점이 있기 때문에, 사용자에게 센서를 장착하여 인식하는 방법이 널리 연구되고 있다. 이를 위해 마이크^[1], 가속도계^[2], 카메라

등 다양한 종류의 센서가 사용되고 있으나, 장소 인식에 있어서는 카메라로부터 얻어지는 영상 신호가 가장 적합한 정보를 담고 있기 때문에 가장 효과적이다.

이러한 이유로, 본 연구에서도 카메라를 이용한 장소 인식 시스템의 개발을 목적으로 한다. 즉, 사용자가 카메라를 장착하고 학습된 환경 내의 여러 장소들을 자유롭게 이동하는 상황에서, 사용자에게 장착된 카메라로부터 연속적으로 입력되는 영상 정보를 분석하여, 사용자의 현재 위치를 실시간으로 인식하여 결과로 내어주는 시스템이다.

그러나 영상은 많은 양의 정보를 담고 있지만 카메라의 흔들림이나 각도, 조명 상태에 따라서 많은 변이가 존재하기 때문에, 영상으로부터 장소를 인식하기 위해서는 다양한 변이에 가능한 인식 방법의 개발이 필수적이다. 이러한 이유로 효율적인 인식 방법의 개발을 위한 연구가 여러 연구 그룹에서 활발히 이루어지고 있다. 기존의 연구된 방법들의 특징을 요약하면 다음과 같다.

먼저 대부분의 기존 연구에서는 정면을 향한 한 대의 카메라만을 사용하였다^{[3][4][5]}. 그러나 이렇게 한 대의 카메라만을 이용한 경우에는 정면 방향의 영상이 장소를 구분하기에 충분한 정보를 담고 있지 않을 경우 정확한 인식이 어려워지는 단점이 있다. 정면 방향의 영상이 현재 위치한 장소에서만 관찰될 수 있는 특징적인 장면이 아닌, 다양한 장소에서 비슷하게 관찰될 수 있는 일반적인 장면을 담고 있거나,

움직임이나 조명 변화 등에 의해 영상에 심한 변이가 발생했을 경우에는 올바르게 장소를 인식하는 것이 어렵다.

다음으로 많은 연구에서 장소간의 시간적인 추론을 하였다^{[3][4]}. 시간적인 추론을 위하여 다음 장소로의 이동 확률이 현재 장소에 의해서만 결정되는 1차 은닉 마르코프 모델을 사용하였다. 이를 통해서 이전 장소와 현재 장소와의 인접 관계와 같은 실제 장소간의 위치관계를 반영할 수 있었지만, 그럼에도 불구하고 다양한 장소간 위치 관계의 일부분밖에 사용하지 못한다는 단점이 있었다.

그 밖의 특징적인 연구 방법으로는 물체 인식을 이용한 장소 인식 방법이 있었다^{[5][6]}. 이는 미리 다양한 물체들에 대한 모델을 학습시키고 영상으로부터 물체 인식을 수행하여, 그로부터 얻어진 물체 인식 결과를 장소 인식을 위한 사전 지식으로 사용하는 방법이다. 물체 정보와 같은 고수준의 정보를 함께 사용함으로써 향상된 인식 성능을 보였으나, 영상으로부터 물체를 인식하기 위해 소요되는 시간이 길어서 실시간으로 사용자의 위치를 인식하는 시스템에 적용하는 것은 실질적으로 불가능하다는 한계가 있었다.

본 연구에서는 카메라 기반 장소 인식을 위해 다음과 같은 방법을 제안한다. 먼저 다중 카메라를 사용하여 여러 방향의 영상 정보를 얻고, 그 정보를 인식에 활용하는 방법이다. 이를 통해 한 방향의 영상 정보만으로는 장소를 구분하기 어려운 경우에도, 다른 방향의 영상 정보를 사용함으로써 보다 정확하게 장소를 인식하는 것이 가능하다. 이를 위하여 여러 대의 카메라를 사용하게 되는데, 많은 수의 카메라를 사용하면 얻을 수 있는 정보량은 증가하지만, 사용자의 자연스러운 행동에는 방해줄 가능성이 커지게 된다. 본 연구에서는 4대의 카메라를 이용하여 전후좌우 네 방향의 영상 정보를 사용할 경우, 그보다 많은 수의 카메라를 이용할 경우에 비해 크게 차이가 없는 충분한 정보를 얻을 수 있으면서도, 사용자의 자연스러운 행동에도 지장을 주지 않는다고 가정한다.

두번째로는, 기존에 사용되던 시간적 추론 방법을 더욱 발전시켜, 다음 장소를 추론하는 데에 현재 장소만을 고려하는 것이 아니라, 현재 장소에서 머무른 시간과 이전 장소를 함께 고려하는 방법을 제안한다. 현재 장소만을 이용할 경우에는 건물 밖에서 한 번에 건물 내부의 회의실로 이동할 수 없다는 식의, 현재 장소와 다음 장소의 인접 관계에 기반한 이동 가능 여부 정도만을 추론에 고려할 수 있다. 그러나 머무른 시간을 함께 고려하므로써, 복도 끝에 위치한 장소로부터 복도의 반대편 끝에 위치한 장소로 이동하기 위해서는 복도에서 일정 시간 이상 머물러야한다는 식의, 장소간의 거리 관계에 기반한 이동 가능 여부도 추론에 반영할 수 있다. 이처럼 다양한 장소간의 위치관계를 반영하여 추론함으로써 보다 사실적인 추론이 가능하다.

2. 장소 인식 시스템

장소 인식 시스템은 현재 시간까지의 영상 정보로부터, 현

재 사용자가 위치해 있을 확률이 가장 높은 장소를 결과로 내어주는 시스템이라 할 수 있다. 이는 사용자의 장소를 내부 상태(Hidden state)로, 영상 신호를 관측값(Observation)으로 하는 은닉 마르코프 모델을 이용하여 모델링 할 수 있다. 시간 t 에서의 사용자의 장소를 Q_t 로, 영상의 특징 벡터를 z_t 로 표기할 때, 주어진 영상 정보들에 대한 현재 각각에 장소에 위치할 확률 $P(Q_t=q|z_{1:t})$ 은 다음과 같은 식으로 계산될 수 있다.

$$P(Q_t=q|z_{1:t}) \propto p(z_t|Q_t=q)P(Q_t=q|z_{1:t-1})$$

where (1)

$$P(Q_t=q|z_{1:t-1}) = \sum P(Q_t=q|Q_{t-1})P(Q_{t-1}|z_{1:t-1})$$

이 식에서, $p(z_t|Q_t=q)$ 는 주어진 장소에서 특정 영상이 관측될 관측확률을 의미하고, $P(Q_t=q|Q_{t-1})$ 는 이전 시간까지의 사용자의 장소로부터, 현재 장소로 이동할 장소이동확률을 의미한다. 장소이동확률은 마르코프 가정을 통해 현재 장소에 영향을 주는 이전 장소의 개수를 제한함으로써 보다 간단하게 쓰여질 수 있다. 그러한 경우 $P(Q_{t-1}|z_{1:t-1})$ 도 장소이동확률과 관측확률을 이용하여 재귀적으로 계산할 수 있으므로, 관측확률과 장소이동확률만을 얻어내면 원하는 확률계산이 가능하다.

2.1장에서는 표본 데이터로부터 관측확률을 구하는 방법을, 2.2장에서는 장소간의 위치관계로부터 장소이동확률을 구하는 방법을 설명하고, 2.3장에서는 구해진 관측확률과 장소이동확률을 이용하여 효율적으로 위의 식을 계산하는 방법을 설명한다.

2.1 관측 확률

2.1.1 특징 추출

영상으로부터 전역적 특징을 추출하기 위해서 다중 steerable pyramid^[7]방법을 사용하였다. 한 영상에 대해 4방향과 4스케일의 steerable pyramid를 적용하여 그림 2와 같은 16개의 분해된 영상을 얻었다. 각각에 영상에 대해 공간적 정보를 유지하는 특징을 추출하기 위하여 영상의 넓은 공간 영역에서 평균된 등급별 중간값을 사용하였다. 즉, 각각의 영상을 4x4 구간으로 나누고 각 구간의 속하는 점들의 평균값을 취하는 방법으로, 영상당 16개, 총 16*16=256개의 특징을 추출하였다. 이렇게 구한 256차원의 특징 벡터를 PCA를 이용하여 80차원으로 축소하여 사용하였다.

2.1.2 관측 확률 추정

각 방향의 영상으로부터 추출된 특징을 이용하여 4방향 영상에 대한 관측확률을 추정하는 것은, 각각의 특징 벡터를 하나로 결합하여 4방향 영상 전체에 대한 특징 벡터를 얻어서 추정하는 방법으로 가능하다. 그러나 현실적으로 전체 영

상에 대한 특징으로부터 확률 분포를 한번에 추정하는 것은 특징 벡터의 차원이 높고 변이가 심하기 때문에 정확한 확률 분포 추정이 어렵다. 따라서 각 방향의 영상이 서로 독립이라는 가정을 두어, 각 방향 영상에 대한 관측확률을 구하고 그것들을 결합하여 전체 영상에 대한 관측확률을 구하는 방법을 사용하였다

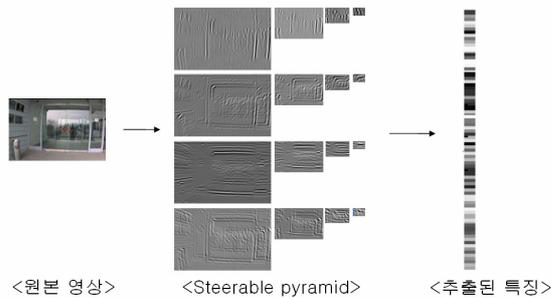


그림 1 Steerable pyramid를 이용한 특징 추출

각 방향 영상에 대한 관측확률 $p(z|Q)$ 은 표본 데이터로부터의 학습을 통해서 추정되는데, 이를 위해서 널리 쓰이는 비모수적 확률 분포 추정 방법인 Parzen window 방법을 사용하였다. Parzen window에 사용되는 윈도우 함수는 일반적으로 흔히 쓰이는 가우시안 함수를 사용하였고, 윈도우의 크기는 실험을 통해 최적값을 구해서 사용하였다.

4방향 영상 전체에 대한 관측확률은 위에서 추정한 각 방향 영상의 관측확률을 다음과 같이 결합하여 구한다.

$$p(z_{all}|Q) = p(z_{front}|Q) \times p(z_{left}|Q) \times p(z_{right}|Q) \times p(z_{back}|Q) \quad (2)$$

2.2 장소이동확률

장소이동확률은 일반적인 은닉 마르코프 모델에서의 전이 확률에 해당한다. 이러한 전이확률에 대한 사전 지식이 없는 경우, 표본 데이터로부터의 학습을 통하여 전이확률을 구하는 것이 일반적이지만, 장소 인식 문제에서는 장소간의 위치 관계로부터 이전 장소에서 이동할 수 있는 장소와 이동할 수 없는 장소를 결정할 수 있다. 따라서 장소이동확률을 표본 데이터를 통해 학습시키지 않고, 이미 알고 있는 장소간의 위치관계를 이용하여 확률값을 결정하는 방법으로 더욱 정확한 모델을 만들 수 있다.

본 연구에서는 다음 장소로의 장소이동확률을 이전 시간까지의 장소 이동 경로에서 다음 장소로의 이동이 적합한지 아닌지에 따라 이전 시간까지의 장소 이동 경로에서 이동 불가능한 장소의 경우에는 0의 확률값을 갖도록 하였고, 이동 가능한 장소에 대해서는 모든 장소에 대해 동일한 확률값을

갖도록 설정하였다. 그리고 다른 장소로 이동하지 않고 현재 장소에서 계속 머무를 경우에 대해서만 일정한 높은 확률값을 갖도록 설정하였다.

이동 가능한 장소에 대해서 모두 동일한 확률값을 갖도록 하는 대신, 이동 가능한 장소에 대해서만 표본 데이터로부터의 학습을 통하여 확률값을 추정하는 것도 가능하지만, 다음과 같은 이유로 사용하지 않았다. 장소 이동 경로는 실제 사용자에게 따라서 주로 이동하는 경로가 다르기 때문에, 다른 사용자들의 데이터를 통해서 학습된 장소이동확률이 실제 사용자에게 대한 값과 다를 가능성이 높다. 이처럼 사용자에게 따라서 장소이동확률이 크게 달라지기 때문에 일반적인 장소이동확률을 학습을 통해 추정하는 것에는 큰 의미가 없다. 어느 사용자에게 대해서도 똑같이 적용되는 가능한 이동 경로와 불가능한 이동 경로에 대한 모델을 구축하는 것이 보다 중요하다.

다음 장소로의 이동 가능 여부를 판단하기 위해서 다음과 같은 두 가지 제약 조건을 사용하였다. 고려할 첫 번째 제약 조건은 장소간의 인접 관계에 기반한다. 이는 사용자가 한 장소에서 다른 장소로 이동하는 것은 반드시 인접한 장소에서만 가능하다는 것으로 이를 ‘인접 장소 제약’으로 명칭한다.

두 번째 제약 조건은 현재 장소에서의 체류시간에 기반한다. 이는 사용자가 현재 장소에서 다른 장소로 이동하기 위해서는 현재 장소에서 일정 시간 이상의 머무름이 필요하다는 것으로 이를 ‘체류시간 제약’으로 명칭한다.

2.2.1 인접 장소 제약

장소 이동에 있어서, 이전 시간까지 강의실에 위치해 있던 사용자가 다음 시간에 건물 옥상에 위치하는 것은 불가능하다. 이는 장소 이동에는 공간적인 제약이 따르기 때문이다. 즉, 공간적으로 떨어져 있는 장소로 바로 이동하는 것은 불가능하고, 사용자는 오직 현재 장소에 계속 머무르거나 또는 인접해 있는 장소로만 이동할 수 있다.

이와 같은 제약을 장소이동확률을 만드는데 적용하기 위해서는, 실제로 어떤 장소들이 인접해 있는지에 대한 정보가 필요하다. 이와 같은 인접 여부를 표현하기 위해 실제 지도로부터 장소 이동 그래프를 구축하여 사용한다. 장소 이동 그래프는 각각의 장소들을 정점(vertex)로 하고, 인접한 장소들에 대해서만 간선(edge)이 존재하는 무방향 그래프(undirected graph)로 표현할 수 있다.

그림 2는 이러한 방법을 통한 추론의 예를 보여준다. 로비, 복도, 연구실, 화장실로 구성된 지도(그림 3(a))로부터 로비와 복도, 복도와 연구실, 복도와 화장실의 인접관계를 표현하는 인접 장소 그래프를 구축하였다(그림 3(b)). 구축한 인접 장소 그래프를 이용하여 이전 시간에는 연구실에 위치해 있었다고 했을 때 현재 시간에서는 복도에 있는지 화장실에 있는지 구분이 어려운 경우, 연구실과 복도는 인접했으므로 이동 가능하지만 화장실은 그렇지 않으므로, 현재의 위치는

복도일 것이라고 추론하였다 (그림 3(c)).

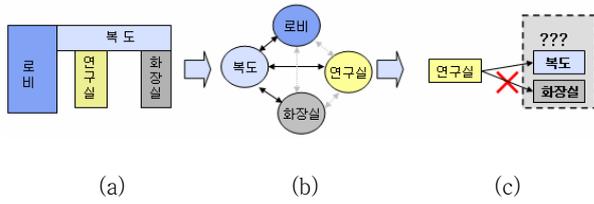


그림 2 인접장소 제약을 통한 추론

인접장소 제약을 통해 다음 장소로의 이동 가능성을 결정하는 것은 현재 장소와 다음 장소와의 관계에만 영향을 받고, 현재 장소에 도착할 때까지 이동한 그 이전 경로에는 영향을 받지 않는다. 따라서 인접장소 제약만을 사용하여 장소이동 확률을 정의할 경우에는 1차 마르코프 모델만으로 다음과 같이 간단하게 표현할 수 있다.

$$P(Q_t=q_j / Q_{1:t}) = P(Q_t=q_j / Q_{t-1}) \quad (3)$$

2.2.2 체류시간 제약

공간적으로 떨어져 있는 장소를 곧바로 이동할 수 없는 것처럼, 멀리 떨어져 있는 장소를 짧은 시간에 이동하는 것도 역시 불가능하다. 즉, 이전 시간까지 1층에 위치해 있던 사용자가 다음 시간에 계단에 위치하고 곧바로 그 다음 시간에 2층에 위치하는 것은, 이동 경로만으로는 ‘1층→계단→2층’의 이동 경로로 생각할 수 있다. 인접 장소 제약만을 고려할 때, 이러한 경로는 인접한 장소로만 이동하여 나타난 경로이기 때문에 적합한 이동 경로로 생각하게 된다. 그러나 현실 상황에서 1층에서 2층으로 이동하기 위해서는 계단에서 어느 정도의 시간이 소요되는데, 이 상황에서는 그러한 시간이 나타나지 않았으므로 현실적으로는 불가능한 상황이다. 이와 마찬가지로 사용자가 복도에서 강의실로 들어갔다 곧바로 복도로 나오는 것과 같이, 한 장소로 이동하자마자 곧바로 원래 장소로 돌아오는 경우도 현실적으로는 거의 일어나지 않는 경우이므로, 이러한 이동 경로도 적합하지 않은 이동 경로라고 가정할 수 있다.

이러한 제약 조건을 정리하면, 한 장소로 이동한 후 곧바로 다음 장소로 이동하는 것은 불가능하다고 할 수 있고, 이를 다시 표현하면 현재 장소에서 다음 장소로 이동하는 것은 현재 장소에서 일정 시간 이상 머무른 이후에만 가능하다고 할 수 있다. 이 때, 이동 경로에 따라서 현재 장소에서 머물러야 하는 시간이 달라지는데, 이러한 장소 이동에 필요한 최소한의 체류시간을 이동 경로별로 미리 구해 놓음으로써, 실제 장소 이동시에 이동 가능 여부를 현재 장소에서 머무른 시간과 이동을 위해 필요한 체류시간을 비교하여 판단할 수 있다.

앞의 예와 같이 현재 장소에서 다음 장소로 이동하기 위

해 필요한 최소 체류시간은 이전 장소와 다음 장소간의 거리에 의해 결정된다. 이전 장소와 다음 장소가 멀리 떨어져 있는 경우에는 그 사이의 장소에서 머물러야 하는 시간이 길고, 가까운 경우에는 짧다. 따라서 각각의 가능한 모든 ‘이전 장소→현재 장소→다음 장소’의 이동 경로에 대해서 현재 장소에서 최소 체류시간이 필요하다.

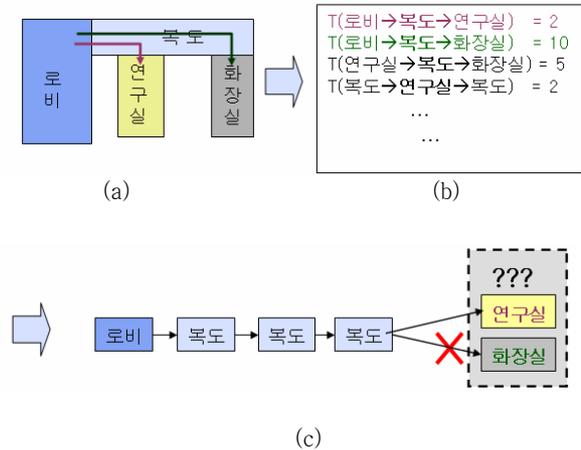


그림 3 체류시간 제약을 통한 추론

그림 3는 체류시간 제약을 이용한 추론의 예를 보여준다. 그림 3(a)와 같은 지도상의 위치 관계로부터 그림 3(b)와 같이 체류시간을 결정하였다. 로비에서 복도를 거쳐 복도 중간에 있는 연구실로 갈때에 비해 로비에서 복도를 거쳐 복도 반대편 끝에 있는 화장실로 갈 때의 복도에서의 최소 체류시간이 더 긴 것과 같이, 이전 장소, 현재 장소, 다음 장소에 따라 결정되는 최소 체류시간 제약이 현실 상황을 잘 표현함을 알 수 있다. 이러한 체류시간을 이용하여 영상 정보만으로는 연구실과 화장실을 구분하기 힘든 경우에, 로비에서 복도로 이동하였고 복도에서 머무른 시간이 3인 상황이라면 이는 ‘로비→복도→화장실’로 이동하기 위한 체류시간 10보다 짧으므로 현재 장소는 화장실이 될 수 없지만, ‘로비→복도→연구실’로 이동하기 위한 체류시간 2보다 길기 때문에 현재 장소는 연구실일 것이라고 추론하였다 (그림 3(c)).

체류시간 제약을 확률 모델로 표현하는 것은 인접장소 제약에 비해 복잡하다. 앞서 설명한 1차 은닉 마르코프 모델로는 이러한 체류시간 제약을 표현할 수 없다. 그 이유는 인접장소 제약에서는 다음 장소로 이동 가능 여부가 현재 장소에 의해서만 결정되는 것에 비하여, 체류시간 제약에서는 현재 장소뿐만 아니라, 이전 장소와 현재장소의 체류시간에도 의존하기 때문이다. 따라서 이러한 관계를 표현할 수 있는 더 복잡한 구조의 모델을 필요로 한다.

다음 상태로의 전이확률에 현재 상태에서의 체류시간을 고려한 모델로 Nonstationary 은닉 마르코프 모델^[8]이 있다. 이 모델에서는 다음과 같이 전이확률을 표현한다.

$$P(Q_t=q_j/Q_{1:t})=P(Q_t=q_j/Q_{t-1}=\dots=Q_{t-\tau}=q_i, Q_{t-\tau-1}\neq q_i) \\ =a_{ij}(\tau) \quad (4)$$

즉, 다음 상태로의 이전 확률이 현재 상태와 현재 상태에서의 머무른 시간에 의해 영향을 받는다고 가정한다. 그러나 체류시간 제약에서의 장소이동확률은, 현재 장소와 현재 장소에서 머무른 시간뿐 아니라, 이전 장소에도 영향을 받기 때문에 이를 표현하기 위하여 기존의 Nonstationary 은닉 마르코프 모델을 확장한 2차 Nonstationary 은닉 마르코프 모델을 제안하였다. 2차 Nonstationary 은닉 마르코프 모델에서의 전이확률은 다음과 같이 표현된다.

$$P(Q_t=q_j/Q_{1:t})=P(Q_t=q_j/Q_{t-1}=\dots=Q_{t-\tau}=q_i, Q_{t-\tau-1}=q_k) \\ =a_{kij}(\tau) \quad (5)$$

이 모델을 통해, 장소 q_i 와 q_j 가 서로 인접하지 않을 경우를 0으로 지정함으로써 인접장소 제약을 표현할 수 있고, τ 가 ' $q_k \rightarrow q_i \rightarrow q_j$ '의 이동 경로에서 q_i 에서의 최소 체류시간보다 작을 경우 $a_{kij}(\tau)$ 를 0으로 지정함으로써 체류시간 제약 역시 장소이동확률로 표현할 수 있다.

2.3 확률 계산

앞서 설명한대로 2차 Nonstationary 은닉 마르코프 모델을 이용하여 인접장소 제약과 체류시간 제약에 의한 장소이동확률을 표현하였다. 2차 Nonstationary 은닉 마르코프 모델에서 시간 t 에서 장소 q_i 에 사용자가 위치할 확률은 다음과 같이 표현된다.

$$P(Q_t=q_j/Z_{1:t})=\sum_i \sum_r P(Q_t=\dots=Q_{t-\tau+1}=q_i, Q_{t-\tau}=q_i/Z_{1:t}) \\ =\sum_i \sum_r b_{t:ij}(\tau) \quad (6)$$

즉, 사용자가 장소 q_i 에서 장소 q_j 로 이동해서 시간 t 까지 τ 시간 동안 머물렀을 경우의 확률을 구하고, 이를 모든 j 와 τ 에 대해 더해서 시간 t 에서 사용자가 장소 q_j 에 위치할 확률을 구할 수 있다. 이 확률값은 $b_{t:ij}(\tau)$ 로 표기하는데, 이는 다음과 같이 재귀적으로 구할 수 있다.

$$\tau > 1 \text{ 일 때,} \\ b_{t:ij}(\tau) = ap(z_t/Q_t)a_{ij}(\tau-1)b_{t-1:ij}(\tau-1) \\ \tau = 1 \text{ 일 때,} \\ b_{t:ij}(\tau) = ap(z_t/Q_t)\sum_k \sum_\delta a_{kij}(\delta)b_{t-1:ki}(\delta) \quad (7)$$

그러나 실제로 한 장소에서 머무르는 시간은 제한되어 있지 않으므로 $b_{t:ij}(\tau)$ 에서의 τ 는 최대 t 까지의 값을 갖을 수 있다. 이러한 모든 τ 에 대해서 $b_{t:ij}(\tau)$ 를 계산하는 것은 많은 계산량이 필요하므로 다음과 같이 계산량을 줄여서 사용한다. 정의된 최소 체류시간중 최대값을 T 라 한다면, 한 장소에서의 머무른 시간이 T 이상일 경우에는 인접한 모든 장소로 이동 가능하므로 일정한 장소이동확률을 갖는다. 따라서 머무른 시간이 T 이상일 경우의 확률을 다음 식과 같이 하나로 묶어서 사용할 수 있다.

$$b'_{t:ij} = \sum_{\tau \geq T} b_{t:ij}(\tau) \\ = \sum_{\tau \geq T} ap(z_t/Q_t)a_{ij}(\tau-1)b_{t-1:ij}(\tau-1) \\ = ap(z_t/Q_t)(a_{ij}(T)b'_{t-1:ij} + a_{ij}(\tau-1)b_{t-1:ij}(\tau-1)) \quad (8)$$

이러한 방법을 사용하였을 때의 확률 계산은 다음과 같다.

$$P(Q_t=q_j/Z_{1:t}) = \sum_i \sum_r P(Q_t=\dots=Q_{t-\tau+1}=q_i, Q_{t-\tau}=q_i/Z_{1:t}) \\ = \sum_i \sum_r b_{t:ij}(\tau) \\ = \sum_i (\sum_{\tau < T} b_{t:ij}(\tau) + b'_{t:ij}) \quad (9)$$

$\tau > 1$ 일 때,

$$b_{t:ij}(\tau) = ap(z_t/Q_t)a_{ij}(\tau-1)b_{t-1:ij}(\tau-1) \\ \tau = 1 \text{ 일 때,} \quad (10)$$

$$b_{t:ij}(\tau) = ap(z_t/Q_t)\sum_k \sum_\delta a_{kij}(\delta)b_{t-1:ki}(\delta) \\ = ap(z_t/Q_t)\sum_k (\sum_{\delta < T} a_{kij}(\delta)b_{t-1:ki}(\delta) + a_{kij}(T)b'_{t-1:ki})$$

이 계산의 계산 복잡도는 인식 대상 장소의 수를 N 으로 표기함으로써 다음과 같이 분석할 수 있다. $b'_{t:ij}$ 의 값과 $\tau > 1$ 일 때의 $b_{t:ij}(\tau)$ 의 값을 구하는 것은 $O(1)$ 의 계산만으로 가능하고, $\tau = 1$ 일 때의 $b_{t:ij}(\tau)$ 의 값을 구하는 것은 $O(NT)$ 의 계산을 필요로 한다. 따라서, 모든 τ 에 대한 $b_{t:ij}(\tau)$ 의 값을 구하는 것도 $O(NT) + (T-1) \cdot O(1) = O(NT)$ 가 되므로 $O(NT)$ 의 계산을 통해 가능하다. 모든 장소에 대해 사용자가 위치할 확률을 구하기 위해서는 위의 계산을 모든 i 와 j 에 대해서 수행해야 하므로 총 계산 복잡도는 $O(N^3T)$ 가 된다.

3. 실험 및 결과

3.1 실험 환경

본 논문에서 제안하는 장소 인식 시스템에 대한 성능 실험을 위해 실제 대학 내의 장소들을 대상으로 인식을 수행하였다. 구체적인 인식 대상 장소는 그림 4에서 보여지는 것과 같은 KAIST 전산학동 내에서의 11개 장소로 설정하였다.



그림 4 실험 대상 장소

실험을 위한 데이터를 수집하기 위해서 여러 명의 사용자에게 카메라를 장착하고, 대상 장소들을 이동 경로를 지정하지 않은 채로 자유롭게 이동하도록 하였다. 사용자가 이동하는 동안 장착된 카메라로부터 연속적으로 영상을 촬영하여 여러 개의 영상 데이터 시퀀스를 얻었고, 이를 시스템의 학습과 성능 평가에 사용하였다.

이를 위해 4대의 웹캠과 미니 PC를 이용하여 실험 장비를 구성하였다. 4대의 웹캠을 조끼의 양쪽 어깨에 부착하고, 사용자는 이 조끼를 입고 미니 PC를 소지한 채로 장소를 이동하도록 하였다 (그림5). 이와 같은 소형의 자연스러운 실험 장비를 제작해 사용함으로써, 사용자가 장비를 착용하지 않았을 때와 차이가 없는 자연스러운 상황에서의 영상 데이터를 수집하였다.



그림 5 실험에 사용된 장비

이러한 방법으로 총 6명의 사용자로부터 각 방향당 12,606장, 총 50,424장의 영상 데이터를 수집하였다. 초당 2~3장의 영상이 수집되었고, 수집된 영상의 해상도는 320x240이었다.

수집된 영상 데이터를 사용자 별로 6개의 그룹으로 나누어, 5개의 그룹을 시스템의 학습에 사용하고 나머지 1개의 그룹에 대해 인식을 수행하는 방법으로 성능을 평가하였다. 그리고, 이러한 평가 방법을 학습에 사용하는 그룹과, 인식에 사용하는 그룹을 바꿔서 반복 평가함으로써 성능 평가의 신뢰도를 높였다. 시스템의 학습을 염두에 두고 영상 데이터를 수집했던 사용자 A를 제외한 나머지 5개의 그룹을 각각 번갈아 학습용 데이터로 사용하는 방법으로 총 5번의 평가 과정을 거쳤다. 최종 인식률의 계산은 각각의 평가에서 얻어진 인식률을 평균하여 구하였다.

4방향 영상을 사용한 것이 인식 성능에 미치는 효과를 분석하기 위해서 정면 방향의 영상만을 사용하여 인식하였을 때와 4방향의 영상을 모두 사용하여 인식하였을 때의 인식률을 각각 구하였다. 또한 체류시간 제약이 인식 성능에 미치는 효과를 확인하기 위해서, 시간적 추론을 사용하지 않고 영상에서의 관측확률만으로 인식하는 방법, 인접장소 제약을 반영한 방법, 인접장소 제약과 체류시간 제약을 반영하는 방법으로 각각 인식기를 구현하여 실험하였다.

3.2 실험 결과

사용한 카메라의 수, 반영된 제약 조건의 정도에 따른 인식률의 비교 결과는 다음과 같았다.

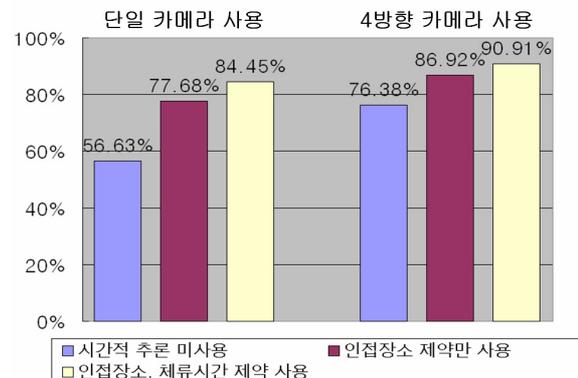


그림 6 인식을 비교 그래프

예상했던 바와 같이, 한대의 카메라를 사용했을 경우 보다 4대의 카메라를 사용했을 경우에 인식률이 상승하는 것을 확인할 수 있었다. 마찬가지로 장소간의 위치 관계를 많이 반영할수록 역시 인식률이 상승하는 것을 볼 수 있었다.

제한하는 연구 방법에 해당하는 다중 카메라와 인접장소 제약, 체류시간 제약을 사용한 인식 방법의 경우 90.91%의 인식률을 나타냈는데, 이는 기존 연구 방법에 해당하는 단일 카메라와 인접 장소 제약만을 사용한 인식 방법의 경우 77.68%의 인식률을 보인 것과 비교하여 13.23%의 인식률 상승을 보였다.

3.3 결과 영상 분석

제안한 방법을 사용함으로써 기존 방법에서 발생했던 오인식이 교정되는 것을 실제 영상 데이터를 통해 분석하였다. 그림 7은 4방향 카메라를 사용함으로써 오류가 없어진 예이다

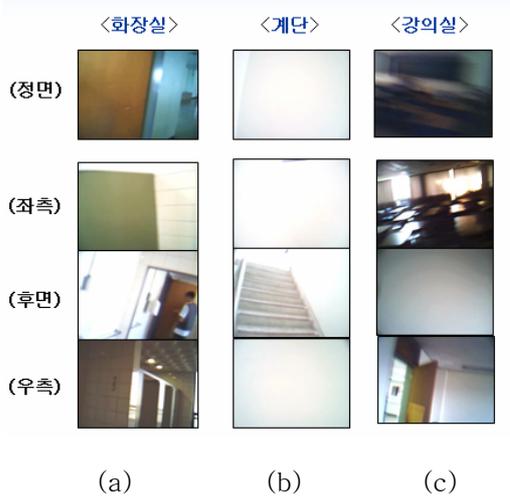


그림 7 다중 카메라 사용을 통해 교정된 오류의 예

그림 7(a)는 실제 사용자가 화장실에 위치하였을 때 촬영된 영상이지만, 정면 방향의 영상만을 사용하여 인식한 경우에는 화장실이 아닌 세미나실로 오인식되었다. 이는 정면 영상에 나타난 장면이 화장실에서만이 아니라 세미나실에서도 나타날 수 있는 장면이었기 때문이다. 그러나 다른 방향에서 관측된 영상에서는 화장실만의 특징적인 장면이 나타났기 때문에 화장실에 대한 확률이 더 높았고, 각 방향의 관측 확률을 결합하여 전체 영상에 대한 확률을 얻은 결과, 화장실이 가장 높은 확률을 얻어서 올바르게 화장실로 인식되었다.

그림 7(b)도 계단에서 촬영된 영상이지만, 사용자가 벽을 바라보고 있을 때 촬영되었기 때문에 정면 영상은 전체가 하얀색으로만 나타났다. 이는 실외나 복도등에서도 사용자가 벽을 바라보고 있었을 상황에서는 흔히 관측될 수 있는 영상이므로 실외로 오인식되었다. 4방향의 영상 정보를 사용했을때에도, 좌측이나 우측의 영상에서는 실외의 확률이 더 높았다. 그러나 그 때의 확률값의 차이는 크지 않았고, 계단부분이 명확하게 관측된 후면 방향의 영상에서 계단이 실외에 비해 크게 높은 확률값을 얻어서 결과적으로 계단으로 올바르게 인식되었다. 그림 7(c)에서도 마찬가지로 정면 영상은 사용자가 이동하는 순간에 촬영되어서 흐려진 영상이 얻어졌고, 그로 인해 사용자가 교실에 위치해있었음에도 계단으로 오인식 되었다. 그러나 4방향의 영상을 모두 고려한 결과 교실로 올바르게 인식되었다.

그림 8은 영상 정보만을 이용하여 장소를 인식한 인식 결과와 시간적 추론을 함께 이용하여 인식한 결과를 비교한 예

이다. 시간적 추론을 사용하지 않고 영상 정보만을 이용하여 인식하였을 때에는, 첫 번째와 두 번째의 영상에 대해서는 사용자의 위치를 연구실로 올바르게 인식하였으나, 세 번째의 영상에 대해서는 로비로 잘못 인식하였다. 그러나 시간적 추론을 사용한 경우에는, 두 번째까지의 영상에서 연구실에 있었을 확률이 가장 높았다는 추론 결과를 이용함으로써 현재 시간에서도 연구실에 위치할 확률이 로비에 위치할 확률보다 높아져서 최종적으로 연구실로 정확히 인식되는 것을 확인할 수 있었다.



그림 8 시간적 추론을 통해 교정된 오류의 예

4. 결론

본 논문에서는 카메라를 이용하여 사용자의 위치를 인식할 때, 현대의 카메라만을 사용함으로써 한 방향의 영상 정보만을 이용하는 것이 아니라, 4대의 카메라를 통해 4방향의 영상 정보를 함께 사용하는 방법을 제안하였다. 이를 통해 단일 방향의 영상만으로는 인식이 어려운 경우에서도, 4방향 영상을 모두 고려함으로써 많은 경우에 대해 옳은 인식결과를 내어줌으로써 전체 인식률이 상승하는 것을 확인할 수 있었다.

또한 기존의 방법에서는 고려하지 못했던 체류시간에 따른 장소 이동 가능 여부를 이용하여 현재 장소를 추론하는 방법과, 확률적 모델에서 이러한 추론을 효율적으로 계산하기 위한 2차 Nonstationary 은닉 마르코프 모델을 제안하였다. 이를 통해 보다 사실적인 추론이 가능해져, 향상된 인식 성능을 얻을 수 있었다

감사의 글

본 연구는 21세기 프론티어 연구개발사업의 일환으로 추진되고 있는 정보통신부의 유비쿼터스컴퓨팅및네트워크원천기반기술개발사업의 지원에 의한 것임

참고문헌

- [1] B.Clarkson, K.Mase, A.Pentland, "Recognizing User Context via Wearable Sensors", Proc. of the 4th IEEE Int. Symposium on Wearable Computers, pp 69~74, 2000
- [2] Seon-Woo Lee, Kenji Mase, "Activity and Location Recognition Using Wearable Sensors", Pervasive Computing, Vol.1, No.3, IEEE, pp. 24~32, 2002
- [3] A. Torralba, K. P. Murphy, W. T. Freeman and M. A. Rubin, "Context-based vision system for place and object recognition", Proc. Of the 9th IEEE Int. Conf. on Computer Vision, 2003
- [4] Fayin Li, J.Kosecka, "Probabilistic Location Recognition using Reduced Feature Set", Proc. of IEEE Int. Conf. on Robotics and Automation, 2006
- [5] S.Kim, I.Kweon, "Model-based Video Interpretation using Static and Temporal Context", Korea-Japan Joint Workshop on Frontiers of Computer Vision, 2006
- [6] 임승빈, 조성배, "동적 베이지안 네트워크를 이용한 컨텍스트 기반 장소 및 물체 인식", 한국컴퓨터종합학술대회 논문집, Vol.33, No.1, pp. 286~288, 2006
- [7] E. P. Simoncelli, W. T. Freeman, "The steerable pyramid: a flexible architecture for multi-scale derivative computation", IEEE Int. Conf. on Image Processing, 1995
- [8] Bongkee Sin, Jin H, Kim, "Nonstationary hidden Markov model", Signal Processing, Vol.46, No.1, pp. 31~46, 1995