

# 패턴비교를 이용한 공연장에서의 관객 수 카운팅 방법에 관한 연구

심 상 균<sup>†</sup> · 박 영 경<sup>\*\*</sup> · 김 중 규<sup>\*\*\*</sup>

## 요 약

본 논문은 극장이나 공연장에 설치된 CCTV 혹은 캠코더를 통해 입력 받은 영상으로부터 패턴비교를 이용해 관객의 수를 카운팅하는 방법을 제안한다. 극장이나 공연장의 환경은 주로 어둡고 관객의 수가 많기 때문에 기존의 객체 검출에 의한 방법으로는 정확한 실시간 카운팅이 어렵다는 문제점을 가지고 있다. 따라서 본 논문에서는 극장이나 공연장의 좌석이 고정되어 있다는 점을 이용해 각 좌석마다 관객이 존재하지 않는 기준영상을 획득한 후 카운팅 하고자 하는 입력 영상의 동일 좌석 영상과 패턴비교를 이용해 각 좌석에서 관객의 유무를 판단한다. 특히, 본 논문에서는 기존 방법들이 지닌 조명 변화이나 잡음에 민감한 문제를 해결하기 위해 상호 보완적인 성능을 보이는 두 가지 특징인 크기투영과 Walsh-Hadamard Kernel에 기반한 패턴비교 방법들을 효과적으로 결합함으로써 조명이나 잡음의 영향을 최소화하고 정확한 카운팅을 가능하게 한다. 제안된 방법은 실제 극장에서 획득한 영상에 대한 실험을 통해 정확한 시스템 성능을 검증하며, 제안된 방법을 주차장에서 빈 주차공간을 확인하는데 적용해 봄으로써 다른 분야로의 적용 가능성을 확인한다.

키워드 : 패턴비교, WHK(Walsh-Hadamard Kernel), 크기투영, 카운팅

## A Study on Audience Counting Method in Auditorium Based on Pattern Comparison

Sang Kyun Sim<sup>†</sup> · Young Kyung Park<sup>\*\*</sup> · Joong-Kyu Kim<sup>\*\*\*</sup>

### ABSTRACT

In this paper, we propose an audience counting method in an auditorium based on pattern comparison. The previous counting methods based on object detection can't exactly count the audience in real time because auditorium has coarse illumination condition and so many audiences. Therefore, in this paper, we count the audience in an auditorium with fixed seats by the method which the pattern from each reference seat is compared to the pattern from each input seat. Especially, to overcome limitations based on either illumination or noise, two pattern comparison methods are efficiently employed and combined. One is based on the amplitude projection, and the other is based on Walsh-Hadamard Kernel. Walsh-Hadamard Kernel has the characteristic which complements amplitude projection. Therefore, we can achieve the accurate counting in the presence of coarse illumination and noise. The experimental results show that our method performs well on sequences of images acquired in an auditorium. We also verify a realistic possibility for other applications applying our method to the parking positioning system.

Key Words : Pattern Comparison, WHK(Walsh-Hadamard Kernel), Amplitude Projection, Counting

### 1. 서 론

최근 감시 및 관측분야에서 영상처리, 컴퓨터 비전, 패턴 인식 등의 기술을 이용하여 제한된 공간에 있는 사람의 수를 자동적으로 측정하는 방법에 관한 연구가 활발하게 이루어지고 있다. 사람의 수를 측정하는 시스템은 극장이나 공

연장에서 관객 수 측정, 지하철 역이나 공항 등에서의 유동 인구 파악, 버스 나 비행기 등의 대중교통 이용객 수 측정 등 다양한 분야에서 폭 넓게 활용될 수 있기 때문에 연구 가치가 크다고 할 수 있다.

현재까지 사람 수 측정에 관한 많은 연구가 이루어져 왔으며 주로 두 가지 유형의 방법들이 제안되어 왔다. 첫 번째는 NeuralNetwork(NN)[1]나 SupportVectorMachine (SVM)[2], AdaptiveBoosting(AdaBoost)[3,4] 등의 객체 학습 방법을 이용하여 사람의 머리 혹은 몸체를 검출하고 이를 기반으로

<sup>†</sup> 정 회 원 : 성균관대학교 대학원 정보통신공학부 석사과정

<sup>\*\*</sup> 준 회 원 : 성균관대학교 대학원 정보통신공학부 박사과정

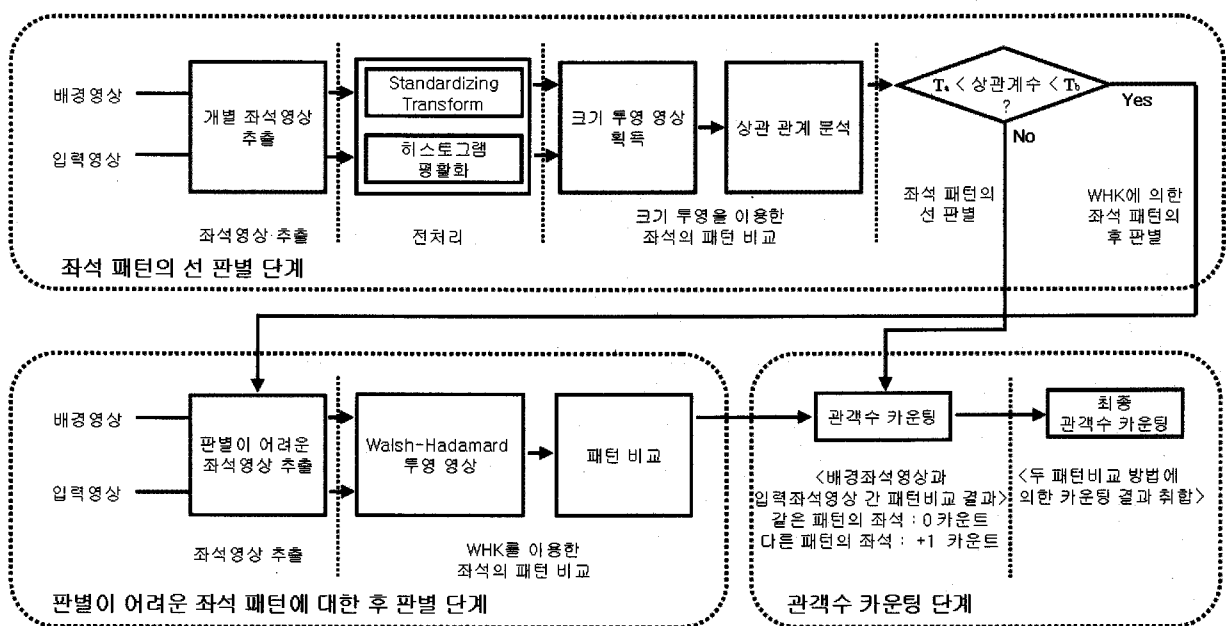
<sup>\*\*\*</sup> 정 회 원 : 성균관대학교 정보통신공학부 교수  
논문접수 : 2006년 9월 28일, 심사완료 : 2007년 1월 2일

사람의 수를 측정 하는 방법이다. 이 방법은 비교적 우수한 성능을 보이지만 다수의 사람이 있는 복잡한 배경이나 저해상도의 영상에서 정확한 객체의 검출이 어려운 단점이 있다. 두 번째 유형은 복잡한 환경에서 사람들이 조밀하게 이동중인 경우 적응적인 배경영상 추출에 의해 사람 그룹을 검출하고 그 수를 추정하는 방법[5,6]으로 복잡한 환경에서는 비교적 강인한 성능을 보이나 그룹 내 사람 수를 정확하게 측정하는 것이 어려운 단점을 가지고 있다. 이러한 두 가지 유형 모두 주로 실내 혹은 실외의 제한된 공간에서 이동중인 사람의 수를 측정하는 데 이용되어 왔다. 지금까지 사람 수 측정에 관한 연구는 주로 제한된 공간에서 이동중인 사람의 수를 측정 혹은 추정하는 방법에 대해 제안되어 온 반면에 적지 않은 수요에도 불구하고 극장이나 공연장에서 관객의 수를 측정하는 방법에 대해서는 아직까지 크게 제안된 바가 없었다. 따라서 본 논문에서는 좌석이 정해져 있고 각 좌석이 관객에 의해 확보되었는지를 판단하는 것이 중요한 극장이나 공연장에서 관객의 수를 측정하는 방법에 대해 제안하고자 한다. 극장이나 공연장은 매우 열악한 조명환경으로 인해 영상의 질이 크게 떨어질 뿐만 아니라 관객의 수가 상당히 많기 때문에 기존의 객체 검출 방법에 의한 사람 수 카운팅으로는 정확한 실시간 카운팅이 불가능하며 사람 그룹의 검출 후 그 수를 추정하는 방법 역시 정확한 카운팅이 힘들다. 또한 두 방법들 모두 각 좌석이 관객에 의해 확보되었는지를 판단하는 것이 어렵다는 문제점을 가지고 있다. 이러한 어려움을 극복하기 위해 본 논문에서는 극장이나 공연장에 배치된 좌석들이 고정되어 있다는 점을 이용한다. 각 좌석마다 관객이 입장하기 전 빈 좌석과 입장한 후 좌석의 패턴을 비교하여 관객이 앉아있는지를 판별한다. 정확한 패턴비교를 위해 본 논문에서는 서로 다른 특성을 보이는 크기투영에 의한 패턴비교 방법[9]과 Walsh-Hadamard Kernel(WHK)[7]에 의한

패턴비교 방법을 이용한다. 먼저 크기투영에 의한 패턴비교 방법을 적용해 관객의 유무를 판별하고 판별 결과가 뚜렷하지 않은 좌석은 WHK에 의한 패턴비교 방법을 이용해 한 번 더 좌석의 패턴을 비교한다. 많은 실험을 통해 크기투영에 의한 방법은 관객이 앉은 좌석에 비해 관객이 없는 빈 좌석을 정확하게 판별하는 반면에 WHK 방법은 빈 좌석에 비해 관객이 앉은 좌석을 정확하게 판별하는 현상을 보였다. 따라서 상호 보완적인 성능을 보인 두 방법을 효과적으로 융합함으로써 단일 방법만을 각각 적용했을 때에 비해 보다 정확한 카운팅이 가능하도록 한다. 또한, 크기투영에 의한 패턴비교 방법을 먼저 적용함으로써 WHK에 의한 패턴비교 방법이 가진 긴 처리 시간의 문제점을 효과적으로 극복할 수 있다.

제안된 방법은 기존의 객체 검출이나 추정을 통해 해결하기 힘든 극장이나 공연장에서의 관객수를 카운팅할 수 있다는 점에서 기존의 방법들과 차별화 되며 두 가지의 패턴비교 방법들을 효과적으로 결합하여 조명이나 잡음 등의 다양한 영향으로 인해 발생할 수 있는 오류를 최소화 할 뿐만 아니라 극장이나 공연장외에도 객체의 자리가 정해져 있는 다양한 장소에 적용하는 것이 가능한데 본 논문에서는 제안된 방법을 주차장에서 빈 주차공간을 확인하는데 적용해봄으로써 이러한 가능성을 보여준다.

본 논문의 구성은 다음과 같다. 먼저 2장에서 본 논문에서 제안한 방법의 흐름도를 보이고, 3장에서 개별 좌석영상의 추출 방법 대해 설명 한 후 4장에서 추출된 개별 영상에 적용된 전 처리 방법을 설명한다. 5장에서는 크기투영과 WHK를 이용한 패턴비교 방법을 이용해 관객의 유무를 판단하고 그 결과를 기반으로 관객수를 카운팅하는 방법에 대해 설명한다. 이 후 6장에서 실험 결과를 보여주고 7장에서는 제안한 방법을 주차장에 적용해 본 후 8장에서 결론을 맺는다.



(그림 1) 제안하는 방법의 흐름도

## 2. 제안하는 관객 수 카운팅 방법의 흐름도

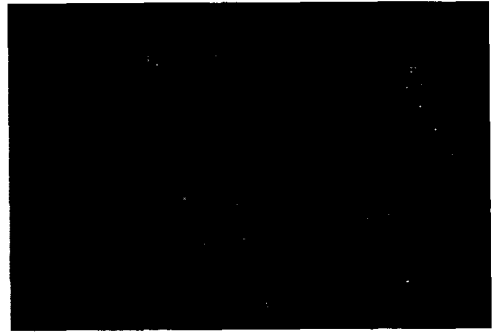
본 논문에서 제안하는 관객 수 카운팅 방법의 구성은 다음과 같다. 먼저 CCTV 혹은 캠코더를 설치하고 먼저 관객이 입장하기 전에 빈 좌석만으로 구성된 배경영상을 획득한 후 관객의 입장이 시작되면 실시간으로 입력영상을 획득한다. 획득된 배경영상과 입력영상에 대해 수동적인 방법으로 각각 개별적인 좌석영상을 추출한다. 배경영상에서 추출된 배경좌석영상과 입력영상에서 추출된 입력좌석영상은 영상의 잡음을 최소화하고 영상의 질을 향상 시키기 위해 전 처리 과정으로써 Standardizing Transform[8]과 히스토그램 평활 화를 순차적으로 적용한다. 전 처리된 배경좌석영상과 입력좌석영상은 픽셀 값의 수직 방향 크기투영 방법을 적용하여 배경좌석영상과 입력좌석영상에 대해 각각 크기투영 영상을 획득하고 획득된 두 영상의 상관관계를 분석하여 좌석의 패턴을 선 판별한다. 그 중 비교적 판별이 어려운 좌석영상은 보다 정확한 패턴비교를 위해 WHK를 이용한 패턴비교를 통해 한번 더 좌석의 패턴을 판별한다. 마지막 카운팅 단계에서는 크기 투영에 의한 패턴비교 결과와 WHK를 이용한 패턴비교 결과를 결합하여 같은 패턴을 보이는 좌석은 빈 좌석이라고 판단하여 +0을 카운트하고 다른 패턴을 보이는 좌석은 관객으로 확보된 좌석이라고 판단하여 +1을 카운트하여 관객 수를 카운팅 한다.

## 3. 좌석의 위치정보에 의한 개별 좌석영상의 추출

각 좌석마다 관객이 입장하기 전과 입장한 후의 패턴을 비교하기 전에 가장 먼저 선행되어야 하는 것은 좌석의 위치를 확보하는 것이다. 따라서 관객이 입장하기 전 획득한 영상에 대해 수동적인 방법으로 각 좌석에 대한 위치정보를 확보하고 이를 기반으로 개별 좌석영상을 추출한다. (그림 2)와 같이 관객이 입장하기 전 획득된 배경영상에서 추출된 좌석영상을 배경좌석영상이라고 정의하고, 관객이 입장한 후 획득된 입력영상에서 추출된 좌석영상을 입력좌석영상이라고 정의한다. (그림 3)은 추출된 개별 좌석영상을 보여준다. 획득된 배경좌석영상과 입력좌석영상은 동일한 위치와 크기를 가지며 패턴비교 시 비교영상으로 사용된다.

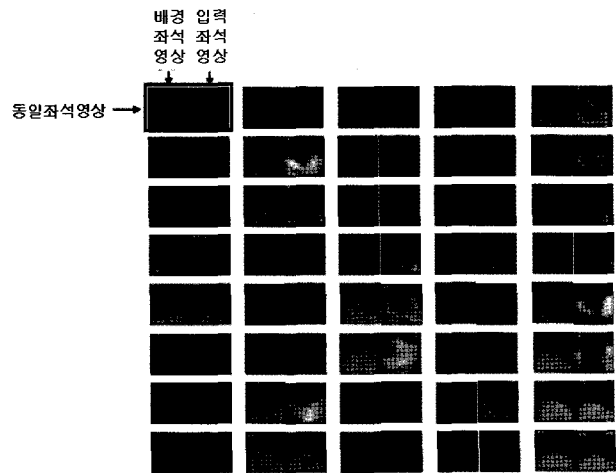


(a) 배경영상



(b) 입력영상

(그림 2) 관객의 입장 전 배경영상과 입장 후 입력영상



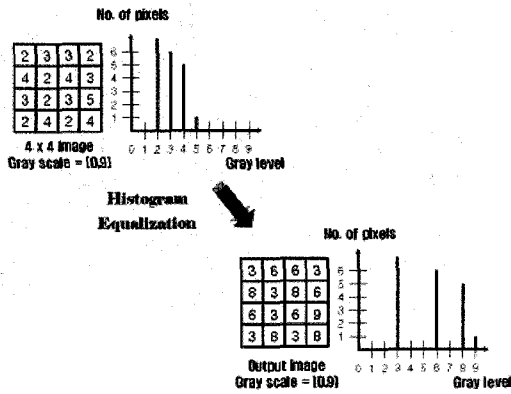
(그림 3) 추출된 개별 좌석영상

## 4. 영상의 질을 향상 시키기 위한 전 처리 과정

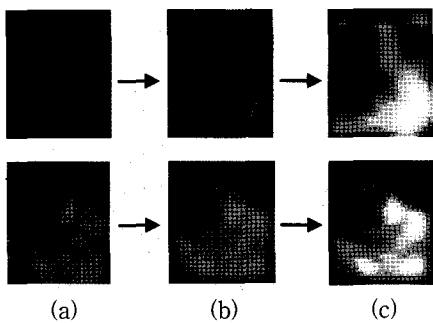
극장이나 공연장 등은 열악한 조명환경으로 인해 영상의 질이 크게 떨어지기 때문에 잡음을 최소화하고 영상의 질을 향상 시키기 위한 전처리 과정은 필수적이다. 본 논문에서는 Standardizing Transform(ST)[8]과 히스토그램 평활 화를 전처리 방법으로 이용한다. ST는 픽셀 간 거리 정보를 이용해 영상을 부드럽게 함으로써 잡음을 최소화시키는 방법이다. 픽셀 간 거리 정보를 나타내는 거리계수( $g_{ij}$ )는 식 (1)을 이용하여 구한다.

$$g_{ij} = f(|P_i - P_j|) = \frac{1}{2\pi\sigma^2} \exp\{-|P_i - P_j|^2 / 2\sigma^2\} \quad (1)$$

여기서  $P$ 는 픽셀을 의미하고  $P_i - P_j$ 는 픽셀  $P_i$ 와  $P_j$  간 거리를 나타내며  $\sigma$ 는 너비 파라미터로서 1로 설정한다. ST처리된 영상은 처리하려는 영상과  $g_{ij}^{1/2}$ 의 곱으로 구할 수 있으며 (그림 5).(b)는 ST처리된 영상을 보여준다. ST에 의해 부드럽게 처리된 영상은 히스토그램 평활 화를 이용해



(그림 4) 영상의 히스토그램 평활화의 예



(그림 5) (a) 원영상 (b) ST 처리된 영상 (c) 히스토그램 평활화 처리된 영상

대비를 향상 시킨다. 히스토그램 평활화는 전반적으로 대조가 작고 영상의 히스토그램이 한쪽으로 치우쳐 있는 영상인 경우에 그 영상의 히스토그램을 넓게 펼쳐줌으로써 영상의 질을 개선하는 것이다(그림 4). (그림 5)(c)는 ST와 히스토그램 평활화를 통해 전 처리된 영상을 보여준다.

**5. 패턴비교에 의한 관객수 카운팅 방법**

관객수 카운팅을 위해 본 논문에서 제안하는 패턴비교 방법은 두 가지 방법으로 구성된다. 먼저 크기투영[9] 영상의 상관계수를 이용해 좌석 내 관객의 유무를 판별한다. 그러나 관객이 좌석의 패턴과 유사한 옷을 입거나 극장 내 조명에 의해 좌석에 그림자가 생길 경우 빠른 연산속도에 비해 비교적 정확성이 떨어지는 크기투영에 의한 패턴비교방법으로는 판별이 힘든 좌석들이 발생한다. 따라서 이러한 분류가 어려운 좌석들은 Yacov Hel-Or이 제안한 Walsh-Hadamard Kernel(WHK)을 이용한 패턴비교[7]를 통해 빈 좌석 또는 관객이 앉은 좌석으로 최종 판별한다. 이 방법은 크기투영에 의한 방법에 비해 처리 시간은 길지만 패턴비교의 대상이 되는 두 영상을 커널 집합에 투영시킨 후 거리 값을 측정하여 패턴을 비교하는 방법으로서 크기투영에 의한 방법이 가진 단점을 보완하여 정확한 패턴의 분류가 가능하다. 본 장에서는 각 방법에 대해 상세히 설명한다.

**5.1 크기투영 영상을 이용한 좌석의 패턴비교**

영상의 질을 향상시키기 위한 전처리 과정을 거친 개별

좌석영상들은 크기투영 영상을 이용하여 좌석의 패턴을 선 판별한다. 크기투영 영상은 수직/수평 방향의 변화 정도를 반영한 영상으로서 원 영상이 가지는 특성을 보여주는데 본 논문에서는 좌석의 특성에 따라 수평 방향에 비해 영상의 변화 정도를 보다 잘 반영하는 수직 방향만을 고려한다. 일반적으로 빈 좌석과 관객에 의해 확보된 좌석은 서로 다른 특성을 갖기 때문에 크기투영 영상을 이용한 패턴비교 방법은 배경좌석영상과 입력좌석영상의 패턴을 비교하기에 적절한 방법이라고 할 수 있다. 또한 이러한 크기투영에 의한 패턴비교 방법은 WHK에 의한 패턴비교 방법에 비해 정확도는 떨어지지만 시간이 적게 걸리며 WHK와는 이용하는 특성에 있어 차이가 있기 때문에 WHK를 이용한 패턴비교 방법에 앞서 적용함으로써 관객수 카운팅 시스템의 속도와 정확도를 효과적으로 개선시킬 수 있다.

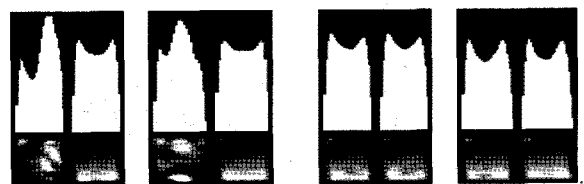
수직 방향 크기투영 영상은 좌석의 픽셀 값을 수직 방향으로 모두 더한 값을 히스토그램 형식으로 나타내며 식(2)와 같이 구한다.

$$V(i) = \sum_{j=1}^n I(i, j) \quad i = 1, 2, \dots, m \quad (2)$$

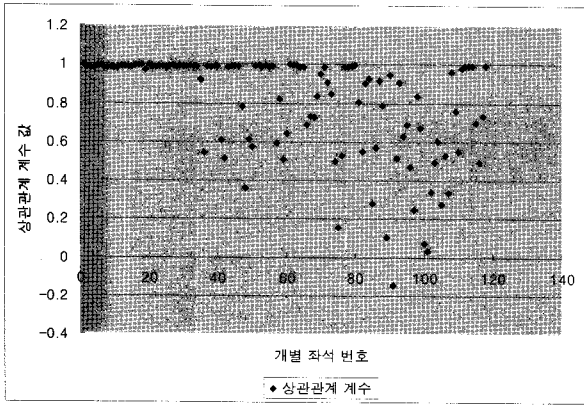
여기서  $i$ 는 영상의  $x$ 좌표이고  $j$ 는 영상의  $y$ 좌표이다.  $m$ 은 영상의 너비( $x$ 축 길이)이고,  $n$ 은 영상의 높이( $y$ 축 길이)이다.  $I(\bullet)$ 는 영상의 픽셀 값을 의미한다. (그림 6)은 빈 좌석과 관객이 앉은 좌석에 대한 수직 방향 크기투영 영상을 보여준다. (a)는 관객이 앉은 좌석에 대한 그림이고, (b)는 빈 좌석에 대한 그림이다. 각 그림의 왼쪽은 입력좌석영상에 대한 원 영상과 크기투영 영상을 보여주고 오른쪽은 배경좌석영상에 대한 원 영상과 크기투영 영상을 보여준다. (그림 6)(a)를 통해 관객이 앉은 좌석의 크기투영 영상은 관객에 의해 배경좌석영상의 크기투영 영상과는 다르게 표현되는 것을 볼 수 있다. 반면 (그림 6)(b)와 같이 관객이 앉은 빈 좌석의 경우 입력좌석영상과 배경좌석영상에서 크기투영 영상이 거의 유사하게 표현 되는 것을 볼 수 있다.

위의 과정을 통해 획득된 배경좌석영상과 입력좌석영상의 크기투영 영상은 식(3)의 상관계수에 의해 빈 좌석과 관객이 앉은 좌석으로 분류한다.

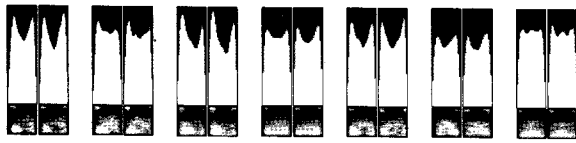
$$\rho_{x,y} = \frac{COV(x,y)}{\sigma_x \sigma_y} = \frac{E[XY] - E[X]E[Y]}{\sigma_x \sigma_y} \quad -1 \leq \rho_{x,y} \leq 1 \quad (3)$$



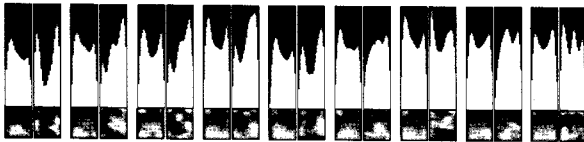
(a) 관객이 앉은 좌석 (b) 빈 좌석  
(그림 6) 수직 방향 크기투영 영상(위)과 원 영상(아래)



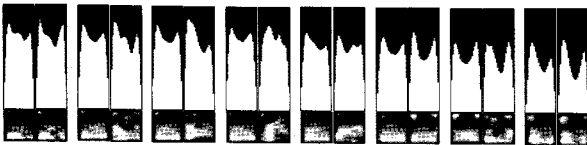
(a) 상관계수( $\rho$ ) 분포



(b)  $\rho > 0.93$



(c)  $\rho < 0.8$



(d)  $0.8 < \rho < 0.93$

(그림 7) (a) 전체 좌석에 대한 상관계수( $\rho$ ) 분포, (b)  $\rho > 0.93$  인 좌석들의 원 영상과 크기투영 영상, (c)  $\rho < 0.8$  인 좌석들의 원 영상과 크기투영 영상, (d)  $0.8 < \rho < 0.93$  인 좌석들의 원 영상과 크기투영 영상

여기서  $X$ 는 배경좌석영상에 대한 크기투영 영상이고  $Y$ 는 입력좌석영상에 대한 크기투영 영상이다.  $COV$ 는 두 영상의 공분산 함수이고,  $\sigma$ 는 표준편차를 의미한다. 두 영상의 상관계수의 절대값이 1에 가까울수록 두 영상이 유사하다고 판단한다. (그림 7).(a)는 모든 좌석에 대한 상관계수의 분포를 보여준다. (그림 7).(a)에서 상관계수가 1에 가까운 좌석들은 (그림 7).(b)와 같이 상관계수가 0.93이상인 빈 좌석들로 왼쪽의 배경좌석영상과 오른쪽의 입력좌석영상의 크기투영 영상이 유사한 형태를 보인다. 이러한 좌석들은 관객이 없는 빈 좌석으로 판단한다. 반면 상관계수가 0.8 이하의 값을 갖는 좌석들은 (그림 7).(c)와 같이 관객이 앉은 좌석들로서 배경좌석영상과 입력좌석영상의 크기투영 영상의 형태가 다른 것을 볼 수 있다. 이러한 좌석들은 관객이 있는 좌석으로 판단한다.

그러나 많은 실험으로부터 좌석에 앉은 사람의 움직임이나 열악한 조명환경 때문에 관객이 앉은 좌석임에도 불구하고

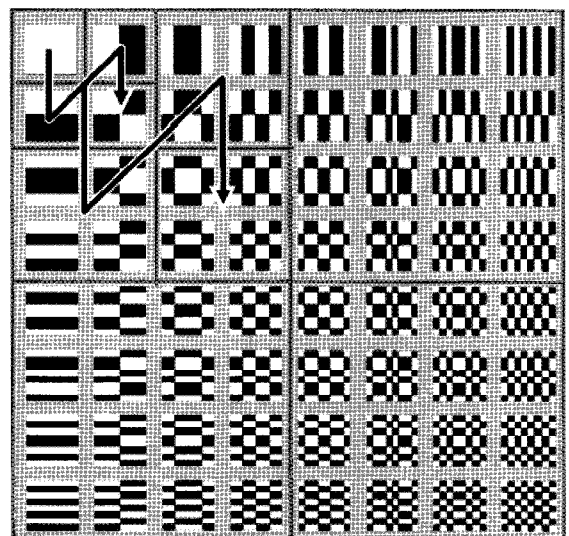
고 배경좌석영상과 입력좌석영상의 크기투영 영상이 유사하게 표현되거나 실제로는 빈 좌석이지만 배경좌석영상과 입력좌석영상의 크기투영 영상이 다르게 표현됨으로써 크기투영에 의한 상관계수로 정확한 판별이 어려운 좌석들이 나타남을 볼 수 있었다. 따라서 (그림 7).(d)와 같이 실험을 통해 설정된 임계범위(0.8~0.93)에 포함되는 좌석들을 판별이 어려운 좌석들로 분류하고 이러한 좌석들은 이후 소개될 WHK를 이용한 패턴비교방법을 이용해 한번 더 좌석의 패턴을 비교한다.

### 5.2 Walsh-Hadamard Kernel을 이용한 좌석의 패턴비교

앞서 소개된 크기 투영영상 기반의 패턴비교방법으로는 배경좌석영상과 입력좌석영상간의 패턴이 유사한 형태를 보이는 좌석에 대해 정확한 판별이 어렵다. 이러한 유사한 패턴의 비교를 위해서 Y.Hel-Or과 H.Hel-Or가 제안한 Walsh-Hadamard Kernel(WHK)에 의한 투영기법에 기반을 둔 패턴비교 방법을 적용한다[7]. Y.Hel-Or과 H.Hel-Or이 제안한 패턴비교 방법은 두 패턴을 패턴비교에 적합한 커널에 각각 투영시킨 후 비교하는 방법인데 Y.Hel-Or과 H.Hel-Or은 투영기법에 적합한 커널로서 Walsh-Hadamard Kernel(WHK)을 선택하여 사용한다[7]. 배경좌석영상과 입력좌석영상을 직접 비교하지 않고 서로 다른 성분을 갖는 여러 개의 커널들에 투영한 후 이들을 비교함으로써 크기투영에 의한 패턴비교방법으로는 판별이 어려운 좌석들에 대한 판별을 가능하게 한다.

#### 5.2.1 트리(Tree)구조를 이용한 WHK 투영기법

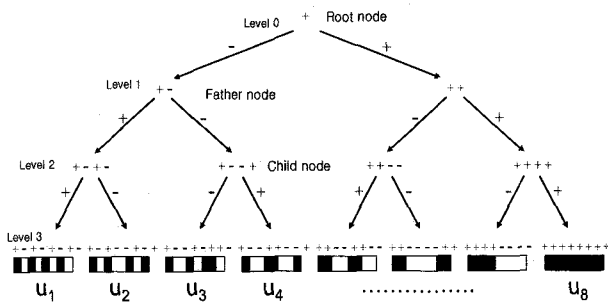
Walsh-Hadamard Transform(WHT)은 Fourier Transform과 더불어 영상의 재 표현 또는 압축 기법으로 아직까지도 많이 사용되는 방법이다. 기저 벡터들이 서로 직교하고 +1과 -1의 이진 값을 갖기 때문에 transform계산 시 덧셈과 뺄셈의 간단한 연산만을 요구하는 장점이 있다. 크기가  $k \times k$  ( $k=2^n$ )인 영



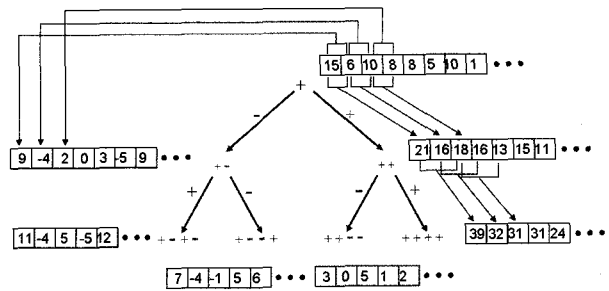
(그림 8) Walsh-Hadamard 투영 벡터집합

상의 WHT는  $k^2$  개의 Walsh-Hadamard(WH) 기저 벡터 집합에 영상을 투영시킴으로써 얻을 수 있다. 이때 각 WH 기저 벡터들은 Walsh-Hadamard Kernel(WHK)을 형성한다. (그림 8)은  $k=8$  인 2차원 WH 투영 벡터집합으로 흰색 직사각형은 +1값을 나타내고 검정색 직사각형은 -1값을 나타낸다. 아래 그림에서 보는 바와 같이 좌측 상단에서부터 수직, 수평방향을 교차하며 이동할수록 투영 벡터의 차수가 증가한다. 투영 벡터의 차수는 패턴비교 시 매우 중요한 역할을 하는데 이는 초기 차수의 투영 벡터에서 두 패턴의 분별력이 높기 때문이다. 예를 들어  $16 \times 16(256)$  크기의 두 패턴을 WH 투영 벡터집합에 투영시킨 후 비교해 보았을 때 전체 투영벡터 256개 중에 초기 10차수 이내의 약 10개의 투영벡터에서 두 패턴의 거리 값의 70% 이상을 갖고 있다. 이것에 대한 설명은 [7]에 자세히 되어 있다.

입력 영상 패턴의 각 WHK로의 투영 값은 (그림 9).(a)와 같이 트리(tree) 구조를 이용하여 효과적으로 계산될 수 있으며 계산 과정은 다음과 같다. 크기가  $k \times k$  인 영상은  $\log_2(k^2)+1$  개의 레벨을 갖는다. 최상위 노드(Root node)는 원 영상(배경좌석영상 혹은 입력좌석영상)이고 최하위  $i$  번째 잎(leaf)인  $U_i$ 는 트리기법을 통해 계산된 원 영상의 WH 투영 벡터이다. 최상위 노드의 원 영상을 포함한 트리 구조 안에 있는 모든 노드의 데이터는 벡터로 표현 되고 상위 레벨 노드의 데이터로부터 계산된다. 각 노드 위의 +(-)는 노드에 있는 데이터에 수행된 모든 연산을 나타내고 상위레벨에서 하위레벨로 연결된 에지 위의 +(-)는 상위 노드의 데이터로부터 하위 노드의 데이터를 구하기 위해 수행된 연산을 나타낸다. 또한 각 노드의 데이터는 상위 노드에 있는 데이터의 모든 원소  $x$ 에 대해  $\Delta$ 거리만큼 떨어져 있는 원소  $x+\Delta$  ( $\Delta=2^{\text{level}}$ )와 에지 위의 기호(덧셈 또는 뺄셈) 연산을 함으로써 구할 수 있다. 예를 들어 크기가 4인 벡터에 대한 WH 투영벡터에 투영한다면 레벨 1에서의 2개의 노드는 레벨 0의 최상위 노드에 있는 4개의 원소에 대해 거리가  $2^0=1$ 인 원소와의 덧셈과 뺄셈연산으로 구할 수 있고 레벨 2에서의 4개의 노드는 레벨 1의 2개의 노드에서 4개의 원소에 대해 거리가  $2^1=2$ 인 원소와의 덧셈과 뺄셈연산으로 구할 수 있다((그림 9).(b)).



(a) 트리 구조에 의한 WHK 투영기법

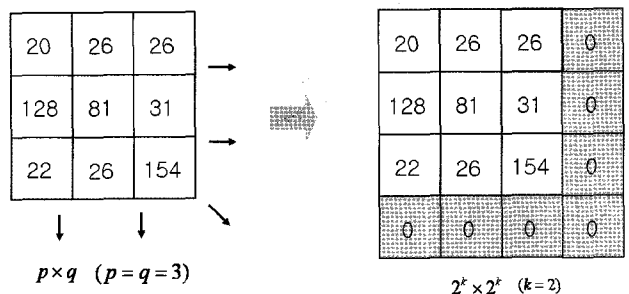


(b) 크기가 4인 벡터에 대한 WHK 투영 계산 과정의 예 (그림 9) 트리 구조에 의한 효율적인 WHK 투영기법

### 5.2.2 투영기법을 이용한 패턴비교

앞서 언급한 바와 같이 두 영상(배경좌석영상, 입력좌석영상)의 패턴은 트리구조에 의해 WHK로 투영된 값들을 이용하여 비교된다. 이때 WHK를 이용한 패턴비교 시 영상의 크기는  $2^k \times 2^k$  픽셀로 제한되는데 본 논문에서는 비교의 대상이 되는 두 영상의 가로와 세로의 크기가 다르기 때문에 (그림 10)과 같이 제로(0) 패딩을 통해 영상의 크기를 정규화 한다. 정규화된 영상은  $16 \times 16$  픽셀 또는  $32 \times 32$  픽셀의 크기를 갖는다. 앞 절에서 설명한 바와 같이 보다 정확하고 빠른 패턴비교를 위해 두 영상을 가장 높은 분별력을 갖는 상위 10개의 WH 투영벡터에만 투영을 하여 패턴을 비교한다. 배경좌석영상을  $R$  이라 하고 입력좌석영상을  $I$  라고 할 때 다음의 방법으로 두 영상을 비교한다. WH 투영벡터 집합은  $\{U_i\}_{i=1}^k$  로 나타낸다. ( $k=10$ )

1.  $R$  과  $I$  를 각각 첫 번째 WH 투영 벡터부터 열 번째 WH 투영벡터까지 투영하여  $R_1 \sim R_{10}$  를 구한다.  
( $R_i = U_i^T R$   $I_i = U_i^T I$ )
2.  $R$  과  $I$  의 Euclidean Distance  $D$  를 구한다.  
( $D = \sqrt{\sum_{i=1}^{10} (R_i - I_i)^2}$ )
3. 거리 값이 실험으로부터 얻은 임계 치  $\alpha$  보다 작다면  $R$  과  $I$  는 유사한 패턴이라고 판별하고, 그렇지 않다면 다른 패턴이라고 판별한다. ( $\alpha = 200$ )



(그림 10) 영상의 크기 정규화를 위한 제로(0)패딩의 예

### 5.3 관객수 카운팅

크기투영과 WHK 투영 방법을 이용해 얻은 결과를 바탕으로 좌석 내 관객의 유무를 결정한다. 두 패턴비교 방법을 적용한 결과 배경좌석영상과 입력좌석영상이 같은 패턴을 보이는 좌석은 관객이 없다고 판단하여 0으로 카운트하고, 다른 패턴을 보이는 좌석은 관객이 있다고 판단하여 1로 카운트한다. 이러한 방식으로 모든 좌석에 대해 카운트를 하여 현재 상영관 내 좌석에 앉은 모든 관객의 수를 카운팅한다. 이때 카운팅 결과의 신뢰도를 높이기 위해 먼저 연속 프레임들에 대해 관객수를 카운팅 한 후 그 결과를 합하여 80%이상 동일하게 나온 카운팅 결과를 최종 카운팅 결과로 결정한다. 관객수를 카운팅을 통해 지정된 영역 내 총 관객의 수를 측정하는 식은 다음과 같다.

$$c = \begin{cases} 0 & \text{같은 패턴을 보이는 좌석} \\ 1 & \text{다른 패턴을 보이는 좌석} \end{cases}$$

(개별 좌석 내 관객의 유/무)

$$C_{k,f_n} = \begin{cases} 0 & \text{if } \sum_{i=f_n}^{f_n+4} c_i \geq 4 \\ 1 & \text{else} \end{cases} \quad (4)$$

(전체 좌석에 대해 카운팅 된 총 관객 수)

$$T = \sum_{k=0}^N C_{k,f_n} \quad (5)$$

$$f_n = 5(n-1) + 1 \text{ where, } n = 1, 2, 3, \dots (\text{frame number}),$$

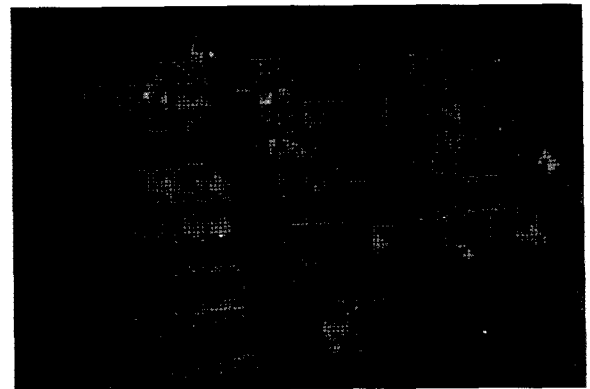
$$k = \text{개별 좌석, } N = \text{총 좌석 개수}$$

## 6. 실험 및 고찰

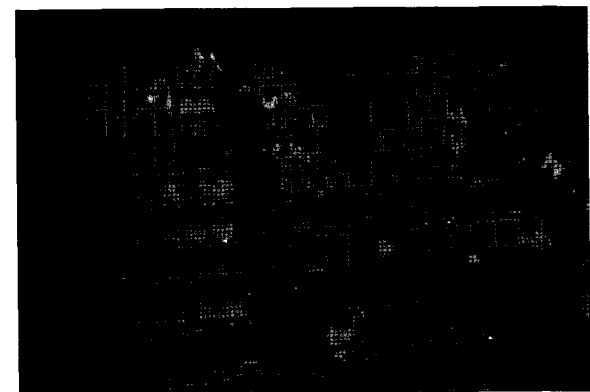
제안된 방법은 Windows XP 운영체제를 갖춘 펜티엄 4 CPU 1.6GHz의 데스크 탑에서 Visual C++ 6.0으로 구현되었다. 실험영상은 300좌석 이상 규모의 상영관에 나이트 샷 기능을 갖춘 SONY DCR-TRV 20캠코더를 설치하여 획득한 720×480 픽셀크기의 비디오 영상에 대해 카메라의 범위에 포함되는 118개의 좌석영상들을 수동적인 방법으로 추출하여 이용하였다. 추출된 개별 좌석영상들은 가로와 세로 모두 15픽셀에서 32픽셀 이하의 다양한 크기를 가지는데 크기투영을 이용한 패턴비교방법에는 원래 크기의 개별 좌석 영상들이 그대로 적용되며 WHK를 이용한 패턴비교 방법에는 5장에서 언급한 바와 같이 영상 크기에 따라 제로(0) 패딩에 의해 16×16픽셀 또는 32×32픽셀의 크기로 정규화된 개별좌석영상들이 적용된다. 제안된 관객수 카운팅 방법의 성능을 측정하기 위해 다양한 실험을 수행하였다. 실험은 크기투영을 이용한 패턴비교 방법과 WHK를 이용한 패턴비교 방법을 개별적으로 적용했을 경우와 두 방법을 융합했을

경우에 대해 이루어졌다. 또한 제안한 방법의 성능은 각 실험에 대한 전체 카운팅 결과와 빈 좌석을 관객으로 확보된 좌석이라고 판단하는 오버 카운팅 결과와 관객으로 확보된 좌석을 빈 좌석이라고 판단하는 미스 카운팅 결과로 측정되었다. 카운팅 결과를 효과적으로 보여주기 위해 실험에 적용되는 모든 좌석의 영역을 빨간색 사각형으로 나타내었으며, 좌석에 관객이 있다고 판단되는 좌석은 초록색 사각형으로 나타내었다.

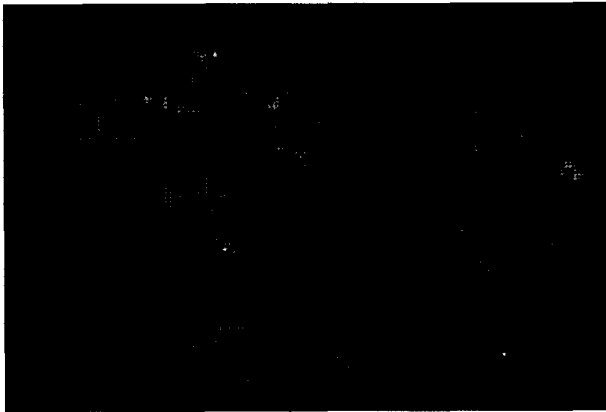
(그림 11)은 동일한 영상에 대해 크기투영 방법과 WHK를 이용한 패턴비교방법을 적용했을 때의 실험결과와 두 방법을 융합했을 때의 실험결과를 보여준다. 그림으로부터 두 방법을 개별적으로 적용했을 때 보다 융합하여 적용했을 때 더욱 정확한 성능을 보임을 알 수 있다. 많은 실험을 통해 크기투영을 이용한 방법이 비교적 관객으로 확보된 좌석에 비해 빈 좌석에 대해서 정확한 판별이 이루어지는 반면에 WHK를 이용한 방법은 빈 좌석에 비해 관객으로 확보된 좌석에 대해서 정확한 판별이 이루어지는 모습을 보였다. 따라서 본 논문에서 제안한 바와 같이 먼저 크기투영을 이용한 패턴비교 방법을 적용하여 빈 좌석이라고 확실히 판단되는 빈 좌석을 먼저 판별한 후 크기투영에 의해 정확한 패턴비교가 어려운 좌석들에 대해 WHK를 이용한 패턴비교 방법을 적용함으로써 좌석 내 관객의 유무에 대한 정확한 판별이 이루어 질것을 쉽게 예측할 수 있다.



(a) 크기투영을 이용한 패턴비교 방법만을 적용했을 경우 ( $\rho < 0.9$ )



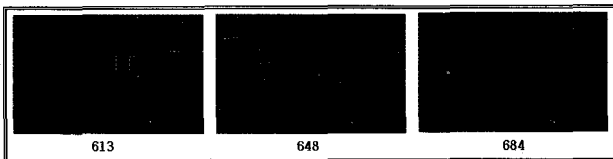
(b) WHK를 이용한 패턴비교 방법만을 적용했을 경우 ( $\alpha = 200$ )



(c) 두 방법을 순차적으로 융합한 경우

(그림 11) 두 방법을 개별적으로 적용했을 경우와 순차적으로 융합하여 적용했을 경우의 관객이 앉은 좌석의 판별 결과

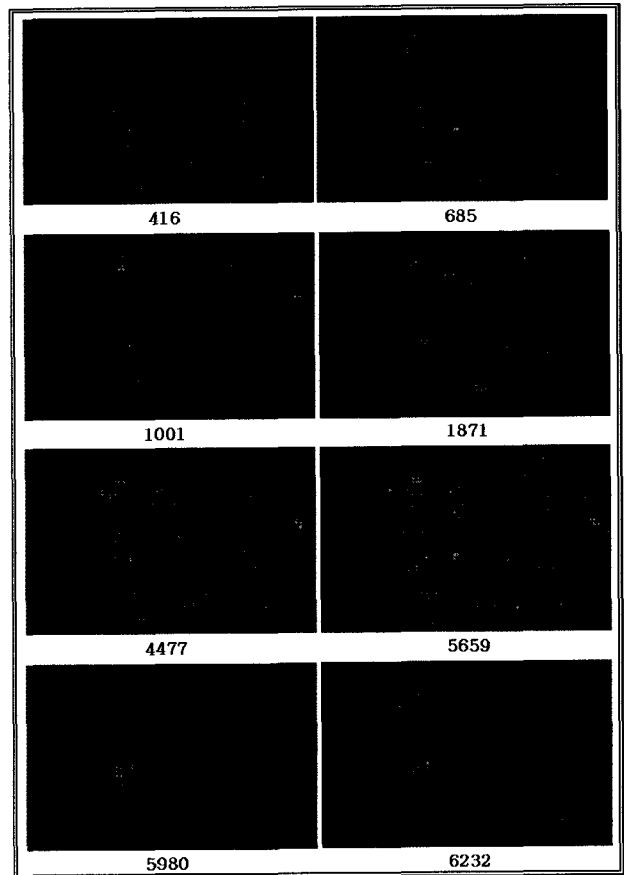
한편 제안된 방법은 (그림 12)와 같이 이동하는 관객에 의해 가려진 빈 좌석을 일시적으로 관객이 앉은 좌석으로 잘못 카운팅하는 오류가 발생한다. 하지만 마지막 관객수 카운팅 단계에서 먼저 연속적인 프레임에 대해 각각 관객수를 카운팅 한 후 80%이상 동일한 결과를 보인 좌석에 대해서만 최종 카운팅을 하는 과정을 통해 이러한 카운팅 오류가 최종 카운팅 결과로 반영되는 현상을 줄일 수 있다.



(그림 12) 좌석으로 이동중인 관객으로 인해 발생한 일시적인 카운팅 오류 현상

(그림 13)은 실제로 관객이 입장하기 전부터 영화가 시작하기 전 극장 조명이 모두 점멸될 때까지 좌석 내 관객의 유무를 판별한 결과를 보여준다. 초기에 관객으로 확보된 좌석의 수가 적을 때나 이후 관객으로 확보된 좌석의 수가 많을 때도 좌석에 대한 판별이 비교적 정확하게 되는 것을 볼 수 있다.

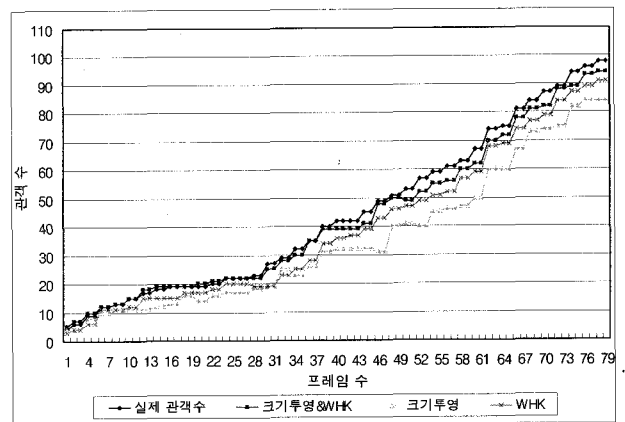
(그림 14)는 실제 상영관에 입장한 관객에 대한 각 방법의 정확도 및 처리 속도에 따른 카운팅 성능의 비교 결과를 보여준다. (그림 14).(a)에서 보는 바와 같이 크기투영에 의한 패턴비교방법과 WHK를 이용한 패턴비교방법을 개별적으로 적용했을 경우 각각 78.04%와 85.87%의 정확도를 보였으며 두 방법을 순차적으로 적용했을 경우 94.86%의 카운팅 정확도를 보였다. 또한 관객이 없는 빈 좌석을 관객이 있는 좌석이라고 카운팅 한 오버 카운팅과 관객 앉은 좌석을 빈 좌석이라고 판단한 미스 카운팅에서도 두 방법을 개별적으로 적용했을 때보다 두 방법을 함께 적용했을 때 카운팅 오류가 감소하는 것을 볼 수 있다(그림 14).(b), (그림 14).(c).



(그림 13) 실제 상황에서의 관객이 앉은 좌석의 검출 결과

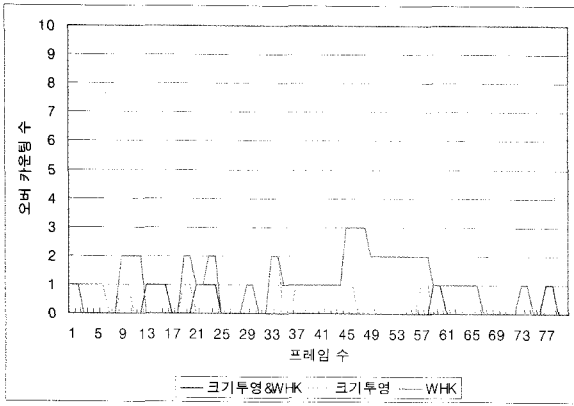
	크기투영 & WHK	WHK	크기투영
카운팅 정확도	94.86%	85.87%	78.04%
처리속도	0.7 fps	0.5 fps	1.4 fps

(a) 각 방법에 대한 평균 카운팅 정확도 및 처리속도 비교

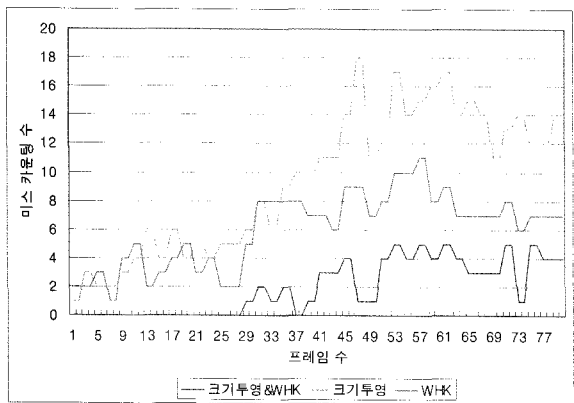


(b) 제안된 방법에 의해 관객수를 카운팅 결과





(c) 오버 카운팅 결과

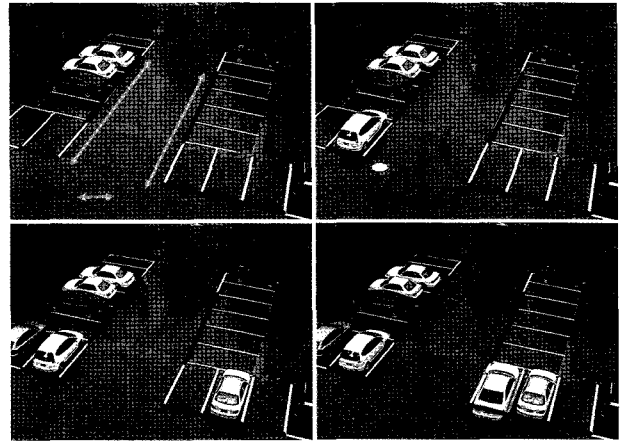


(d) 미스 카운팅 결과

(그림 14) 카운팅 정확도 실험 결과

### 7. 제안한 방법으로 적용 가능한 관련 어플리케이션

본 논문에서 제안한 패턴비교를 이용한 관객수 카운팅 방법은 감시 및 보완분야 중 야외주차장에서 자동차 수 카운팅 및 빈 주차공간 위치 확인시스템이나 건물의 현관 및 출입문에서 들어오는 사람 수 카운팅 시스템 등의 여러 어플리케이션으로 적용이 가능할 것으로 판단된다. 따라서 본 논문에서는 이러한 어플리케이션 중 빈 주차공간의 위치를 확인하는 시스템에 제안한 방법의 적용 가능성을 확인해보기 위해 (그림 15)와 같이 주차장에서 간단한 실험을 해보았다. 각 그림에서 왼쪽의 5자리와 오른쪽의 7자리를 대상으로 실험을 수행한 후 빈 주차공간으로 판단된 자리는 검정색 선으로 표시하였으며 차로 확보되었다고 판단되는 자리는 빨간색 선으로 표시하였는데 (그림 15)에서 보는 바와 같이 정확하게 빈 주차공간을 찾는 모습을 볼 수 있었다. 하지만 실험을 통해 설치되는 카메라의 위치와 각도가 시스템 성능에 중요한 부분을 차지 할 것이라는 예측을 할 수 있었다. 그 이유는 카메라의 위치 및 각도에 따라 주차되어 있는 자동차가 주변의 주차가 가능한 빈 자리를 부분적으로 가리기 때문인데 이러한 현상이 두드러질 경우 카메라의 위치 조정이나 추가적인 카메라 설치가 필요할 것으로 판단된다. 따라서 향후 이러한 부분에 대한 추가적인 연구가 진



(그림 15) 제안한 방법으로 주차장에서 빈 주차공간 확인 실험 결과

행된다면 빈 주차공간 확인 시스템으로 적용이 가능할 것이다. 또한 이러한 야외주차장에서 빈 주차공간의 위치를 확인하는 시스템을 보편화 되어 있는 네비게이션 시스템과 연동하여 아파트 주차장이나 빌딩 주차장 등에 적용할 경우 사용자가 보다 쉽게 빈 주차공간을 확인할 수 있을 것이다.

### 8. 결론

본 논문에서는 크기 투영과 Walsh-Hadamard Kernel을 이용한 좌석의 패턴비교방법에 기반하여 공연장에서 관객의 수를 카운팅하는 방법을 제안하였다. 제안된 방법은 일반적으로 조명환경이 열악해서 기존의 객체 검출에 의한 사람 수 카운팅 방법으로는 관객의 수를 정확하게 카운팅 하는 것이 힘들고, 각 좌석이 관객에 의해 확보되어 있는지를 판단하는 것이 중요한 공연장이나 극장의 특성에 맞도록 패턴 비교방법에 기반하고 있다. 처리시간이 비교적 빠르고 빈 좌석에 대해 비교적 정확한 카운팅 성능을 보인 크기투영에 의한 패턴비교방법을 크기투영에 비해 상대적으로 처리시간은 길지만 관객으로 확보 된 좌석에 대해 정확한 카운팅 성능을 보인 WHK를 이용한 패턴비교방법에 앞서 적용함으로써 두 방법이 가지는 단점을 효과적으로 극복하였다. 또한 연속한 프레임들에 대해 관객수를 카운팅한 뒤 일정횟수 이상 동일하게 나온 결과에 대해 최종 카운팅 결과로 결정하여 카운팅 결과의 신뢰도를 높일 수 있었고 실제로 제안된 방법을 극장환경에서 테스트한 결과 94.86%의 높은 카운팅 정확도를 얻을 수 있었다. 그러나 제안된 방법은 영상 간 패턴비교에 기반하기 때문에 사람이 아닌 옷이나 가방 등의 물건이 좌석에 있을 경우 카운팅 오류가 발생할 가능성이 존재한다. 따라서 향후 이러한 문제점에 대한 보다 많은 테스트를 통해 제안된 방법을 개선해 나가고자 한다.

한편 본 논문에서 제안한 방법은 간단한 실험을 통해 보였듯이 주차장에서 빈 주차 공간을 확인하는 시스템등과 같이 정해진 자리에서 객체의 존재 여부를 판단하는 다양한 어플리케이션으로의 적용이 가능할 것으로 기대된다.

참 고 문 헌

[1] H.A.Rowley, S.Baluja and T.Kanade, "Neural Network-Based Face Detection," IEEE Trans. Pattern Analysis and Machine Intelligence, vol.20, no.1, pp. 23~38, January, 1998.

[2] E.Osuna, R.Freund and F.Girosi, "Training Support Vector Machines: an Application to Face Detection," Proc. CVPR'97, Puerto Rico, June, 1997.

[3] P. Viola, M.Jones, "Rapid object detection using a boosted cascade of simple features," Conference on Computer Vision and Pattern Recognition, 2001.

[4] P. Viola, M.Jones, "Detecting Pedestrians Using Patterns of Motion and Appearance" Proc. IEEE Int. Conf. Computer Vision, Nice, France, Sept., 2003, pp.734-741.

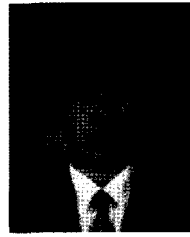
[5] Yongtea Do, "Region Based Detection of Occluded People for the Tracking in Video Image Sequences" CAIP 2005 LNCS 3691, pp 829-836, 2005.

[6] C.Sacchi, G.Gera, L. Marcenaro, and C. S. Regazzoni "Advanced image-processing tools for counting people in tourist site-monitoring applications," Signal Process., vol.81, pp.1017-1040, 2001.

[7] Hel-Or.Y, Hel-Or. H, "Real-Time Pattern Matching Using Projection Kernels" IEEE Trans. Pattern Analysis and Machine Intelligence, Vol.27, No 9, pp 1430-1445, Sept., 2005.

[8] Wang.L, Zhang.Y, Feng.J, "On the Euclidean Distance of Images" IEEE Trans, Pattern Analysis and Machine Intelligence, Vol. 26(2004)4 13.

[9] Huchuan Lu, Youjun Wang, Zunyi Wang, "A Crowd Detecting and Tracking Algorithm Based on Projection of Grayscale Image," Proc. ICIG'04, pp. 188-191, 2004.



심 상 균

e-mail : ssk22ssk@skku.edu  
 2005년 홍익대학교 전자공학과(학사)  
 2005년~현재 성균관대학교 대학원  
 정보통신공학부 석사과정  
 관심분야 : 컴퓨터 비전, 패턴 인식,  
 영상신호처리



박 영 경

e-mail : multipym@skku.edu  
 2002년 성균관대학교 전자공학과(학사)  
 2004년 성균관대학교 대학원 전기전자 및  
 컴퓨터공학부(공학석사)  
 2004년~현재 성균관대학교 대학원  
 정보통신공학부 박사과정  
 관심분야 : 컴퓨터 비전, 패턴 인식, 영상신호처리, 데이터 방송



김 중 규

e-mail : jkkim@skku.edu  
 1980년 서울대학교 전자공학과(공학사)  
 1982년 서울대학교 전자공학과(공학석사)  
 1989년 The University of Michigan  
 Electrical Engineering Ph.D.  
 1980년~1981년 한국통신기술연구소(현  
 ETRI) 위촉연구원  
 1989년~1990년 The University of Michigan, Post Doctoral  
 1990년~1992년 삼성전자 특수연구소 TASR 개발팀장  
 1992년~현재 성균관대학교 정보통신공학부 교수  
 관심분야 : 컴퓨터 비전, 패턴 인식, 의용공학, 적응신호처리,  
 데이터 방송