

화상 인식 서버 기반 감시 카메라 시스템의 성능 분석

쉬키라트 야라*, 이 구 연^o

Performance Analysis of Surveillance Camera System Based on Image Recognition Server

Yara Shqairat*, Goo Yeon Lee^o

요 약

본 논문에서는 화상인식서버 기반의 감시 카메라 네트워크 시스템에서 프레임 폐기율 및 서버의 활용도에 기반한 성능분석을 수행한다. 각 감시 카메라에 대하여 인식상태 및 유희상태로 나누어 다른 프레임 전송률을 적용함으로써 서버의 처리용량을 기준으로 한 최적의 화상 프레임 수 및 수용 가능한 카메라 수 등의 파라미터들에 대하여 분석한다. 분석된 결과는 화상 감시 카메라 네트워크의 효율적인 운영을 위하여 유용하게 활용될 것으로 판단된다.

Key Words : Surveillance, Camara, Image, Recognition, Network

ABSTRACT

In this paper, we perform a performance analysis of a surveillance camera network system with an image recognition server based on frame discard rate and server utilization. Surveillance camera states are divided into recognition states and silence states to analyze the various parameters such as the optimum number of image frames and the optimum number of cameras based on the processing capacity of the server. The analyzed results will be useful for efficient operation of the evolving surveillance camera network systems.

I. 서 론

화상인식서버가 위치하는 중앙감시센터는 감시 카메라들로부터 가능한 많은 화상 프레임을 전달받는 경우 최고 품질의 감시 기능을 수행할 수 있게 된다. 하지만 감시 시스템 운영자는 전송되는 화상 프레임의 수에 따라 네트워크에 적지 않은 비용을 지급해야 하므로 무조건 많은 프레임을 전송하게 할 수는 없게 된다. 이에 본 연구에서는 화상인식서버가 감시카메라가 전송하는 화상 정보에 대하여 화상인식을 수행하여 관심 객체가 있고 없음을 판단하고 이에 따라 각 감시 카메라의 상태를 초당 많은 프레임을 전송하는 인식상태(recognition state)와 그렇지 않은 유희상태(silence state)로 설정하는 시스템을 가정한다. 이와 같은 감시카메라 네트워크 시스템에서 시스템 운영자는 화상인식서버의 처리 용량에 맞추어 네트워크를 통하여 전송되는 화상 프레임의 수를 최적화 할 필요가 있고, 또한 적절한 화상 카메라 수에 대한 설계가 필요하므로, 본 연구에서는 이에 필요한 이론적인 성능 분석 결과를 제시한다. 감시 카메라 시스템에서의 전송 프레임 수나 네트워크 대역폭 및 보안방식에 대한 많은 연구가 수행되어왔다. [1]에서는 감시에 필요한 효과적인 프레임수에 대한 연구가 진행되었으며 [2]에서는 네트워크를 통한 화상 프레임 데이터의 흐름 및 양에 대한 네트워크 및 인프라 설계에 대한 연구가 이루어졌다. 또한 [3]에서는 감시카메라 시스템에서의 보안측면에 대하여 연구하였다.

II. 분석 모델 및 결과

본 연구의 분석 모델에서는 N개의 감시 카메라가 네트워크를 통하여 중앙감시센터의 화상인식서버에 연결되어 있다고 가정한다. 화상인식서버는 도착한 화상 프레임에 대하여 화상 인식을 한 후 관심 객체가 인식된 경우에 해당 카메라를 프레임 전송률이 R_{rec} 인 인식상태로 설정하고, 그렇지 않은 경우에는 프레임 전송률이 R_{sil} 인 유희상태로 설정 한다(단 $R_{rec} > R_{sil}$). 그림 1은 카메라의 인식상태와 유희상태간의 전이과정을 보여준다. 그림 1에서 각 카메라의 인식상태와 유희상태는 번갈아 나타나고 있으며, 이러한 전이과정은 다른 카메라와 독립적으로 이루어진다

* First Author : Kangwon National University, Dept. of Comp. and Comm. Eng., yara_shqairat@hotmail.com, 학생회원

^o Corresponding Author : Kangwon National University, Dept. of Comp. and Comm. Eng., leegyeon@kangwon.ac.kr, 종신회원
논문번호 : KICS2017-02-051, Received February 24, 2017; Revised March 13, 2017; Accepted March 21, 2017

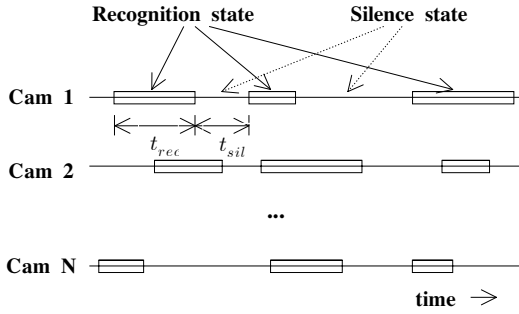


그림 1. 감시 카메라의 인식상태와 유휴상태간의 전이과정
Fig. 1. Transitions between recognition and silence states of surveillance cameras

고 가정한다. 이러한 가정은 감시 카메라가 근거리에서 집중적으로 배치되고 있는 경우에는 적절하지 않으나, 본 연구에서는 넓은 지역에 감시 카메라가 어느 정도 거리를 갖고 분산되어 배치되어 있다고 가정한다. 이 경우에는 각 카메라의 상태전이는 서로 간에 독립적으로 동작한다고 볼 수 있다. 또한 분석의 편의상 모든 카메라의 상태 특성은 동일하다고 가정한다.

그림 1에서 인식상태 구간 길이를 r.v. t_{rec} 로 나타낸다고 가정하고 유휴상태의 구간 길이를 r.v. t_{sil} 로 나타낸다고 가정한다. 또한 화상인식서버의 초당 처리 프레임 수(처리 용량)은 C 라고 가정한다. 일반적으로 감시 시스템은 실시간 감시가 효율적이므로 화상인식 서버는 이 실시간성을 보장할 정도의 약간의 대기큐를 가지고 있다고 가정한다. 이 대기큐의 길이에 따라 지연에 대한 완충 효과가 일부 나타날 수 있으나, 본 연구에서는 분석의 편의상 초단위로 카운트하여 화상인식서버의 처리 용량 C 를 넘어서서 도착하는 프레임은 버려진다고 근사화한다. 위의 가정으로부터 하나의 감시 카메라가 인식상태에 있을 확률 및 유휴상태에 있을 확률은 각각 $P_{rec} = \frac{E(t_{rec})}{E(t_{rec}) + E(t_{sil})}$ 및 $P_{sil} = 1 - P_{rec}$ 으로 구할 수 있다. 여기서 $E(\cdot)$ 는 기대치를 나타낸다. 또한 k 개의 카메라가 인식상태에 있을 확률은 $P_k = \binom{N}{k} (P_{rec})^k (1 - P_{rec})^{N-k}$ (단 $0 \leq k \leq N$)로 구할 수 있다. 운영측면에서 초당 최소 총 전송 프레임수는 화상 인식 서버의 처리 용량 C 보다 같거나 작아야 하며, 또한 C 는 초당 최대 총 전송 프레임수보다 같거나 작아야 한다. 즉 $N \cdot R_{sil} \leq C \leq N \cdot R_{rec}$ 의 관계가 성립하도록 설계해야 한다. 그리고 초당 평균 총 전송 프레임 수는 $N \cdot P_{rec} \cdot R_{rec} + N \cdot (1 - P_{rec}) \cdot R_{sil}$ 와 같이 주어진다. 서버의 처리

용량 C 와 $N \cdot R_{rec}$ 을 같게 설계하는 경우 서버가 최대 용량을 커버하므로 서버에서 처리되지 못하고 버려지는 프레임은 없게 된다. 그러나 이 경우 초당 평균 총 프레임 수에 대하여 평균적으로 초당 $C - N \cdot P_{rec} \cdot R_{rec} - N \cdot (1 - P_{rec}) \cdot R_{sil}$ 의 프레임 수만큼의 서버의 처리용량의 낭비가 있게 된다. 그러므로 서버에서 처리되지 못하고 버려지는 프레임 수를 일부 허용함으로써 서버의 가동률을 높일 필요가 있다. 이에 서버의 처리 용량과 버려지는 프레임의 수, 화상 카메라의 전송률과의 관계에 대한 분석이 필요하다. 서버의 처리용량으로서 평균적으로 처리할 수 있는 정상상태에서의 카메라의 수 m 은 $C = m \cdot R_{rec} + (N - m) \cdot R_{sil}$ 의 식에서 구할 수 있다. 여기서 m 은 정수이어야 하므로 $m = \lfloor \frac{C - N \cdot R_{sil}}{R_{rec} - R_{sil}} \rfloor$ 과 같이 나타낼 수 있다. 단 $\lfloor x \rfloor$ 는 바닥함수로서 x 보다 같거나 작은 정수 중에서 가장 큰 값을 나타낸다. 이 때 서버에서 처리되지 못하고 버려지는 평균 프레임 수의 기대치는 다음과 같이 구해진다.

$$F_{discarded} = \sum_{k=m+1}^N P_k [kR_{rec} + (N-k) \cdot R_{sil} - C] \quad (1)$$

본 연구에서는 성능 척도로서 다음의 두가지를 고려한다. 첫 번째는 서버에 도착했으나 서버에서 처리되지 못하고 버려지는 프레임의 수를 서버의 처리용량 대비로 나타낸 수치 DR (discarded ratio)이다. 이는 버려질 프레임인데도 불구하고 시스템 운영자는 의미없이 네트워크에 비용을 지급하게 된 프레임들로서 이 수치가 클수록 비효율적인 측면이 커진다고 볼 수 있다. 다른 하나는 서버의 활용도로서 서버 용량 C 에 대하여 실제로 서버가 처리하는 프레임의 수의 비율 ρ 로 나타낸다. 이 비율이 높으면 서버의 가동률이 높아 서버에 대한 효율적인 운영이 되는 경우이고, 반대로 이 비율이 낮으면 서버의 가동률이 낮아 주어진 서버 능력에 대하여 비효율적으로 운영하는 경우가 된다. 먼저 DR 은 $\frac{F_{discarded}}{C}$ 로써 주어지며 다음과 같이 나타낼 수 있다.

$$DR = \frac{1}{C} \cdot \sum_{k=m+1}^N P_k [(k \cdot R_{rec} + (N - k) \cdot R_{sil}) - C] \\ = \sum_{k=m+1}^N \binom{N}{k} (P_{rec})^k (1 - P_{rec})^{N-k} \cdot \left[\frac{k \cdot R_{rec} + (N - k) \cdot R_{sil} - C}{C} \right] \quad (2)$$

그리고 ρ 는 다음과 같이 구할 수 있다.

$$\rho = \frac{1}{C} \left[\sum_{k=0}^m P_k \cdot [(k \cdot R_{rec} + (N-k) \cdot R_{sil})] + \sum_{k=m+1}^N P_k \cdot C \right] \\ = 1 - \sum_{k=0}^m \binom{N}{k} (P_{rec})^k (1 - P_{rec})^{N-k} \cdot \left[1 - \frac{k \cdot R_{rec} + (N-k) \cdot R_{sil}}{C} \right] \quad (3)$$

그림 2와 3은 $C=100$, $R_{rec}=5$, $R_{sil}=1$ 일 때 $P_{rec}=0.1, 0.2, 0.3, 0.4, 0.5$ 로 변환에 따른 N 과 DR 그리고 N 과 ρ 의 관계에 대한 수치분석(실선 및 점선) 및 모의실험 결과(\bullet 표시)를 나타낸 그래프이다.

모의실험은 인식상태의 평균시간을 3초로하고, 유휴상태의 평균시간을 각각 27초, 12초, 7초, 4.5초, 3초로 하였으며, 각 상태의 길이는 지수분포를 따르도록 하였다. 그림 2 및 3에서 대기큐 유무의 가정에 의한 미미한 차이를 제외하면 수치분석 결과와 모의실험결과가 거의 일치함을 볼 수 있으며, 이는 본 수치분석 결과가 유효함을 보여준다. 그림 2 및 3에서 $DR=0$ 가 되도록, 즉 서버에서 처리되지 못하고 버려지는

화상 프레임이 하나도 없도록 설계할 경우, 화상 카메라의 수는 $N=20$ 개로 제한된다. 그러나 이 경우 서버의 활용도 ρ 는 $P_{rec}=0.1, 0.2, 0.3, 0.4, 0.5$ 일 때 각각 28%, 36%, 44%, 52%, 60%에 불과하다. 즉 서버의 활용도가 낮아 많은 비효율적인 상황이 발생한다. 서버의 활용도를 높이기 위하여서는 DR 을 어느 정도로 허용해야 한다. $DR=0.01$ (1%)로 하는 경우 $P_{rec}=0.1, 0.2, 0.3, 0.4, 0.5$ 일 때 각각 $N=64, 49, 40, 34, 29$ 의 값을 갖게 되어 최대 설치 가능한 화상 카메라의 수가 커지게 된다. 이때의 서버 활용도도 각각 88.8%, 87.3%, 87.1%, 87.4%, 86.4%가 된다. 즉 서버에서 처리되지 못하고 버려지는 화상 프레임의 수를 서버 처리용량 대비 1%를 허용함으로써 서버의 활용도는 86%~89%로 급격히 증가하게 된다.

III. 결론

본 연구에서는 감시카메라 시스템에서 화상인식서버의 활용도 및 전달 프레임의 폐기율에 기반하여 시스템 설계시에 필요한 여러 성능 요소들간의 관계를 분석하였으며, 이러한 연구 결과는 최근 광범위하게 보급되고 있는 감시 카메라 시스템의 효율적인 설계 및 운영에 유용하게 활용될 것으로 판단된다.

References

- [1] H. Keval and M. Angela Sasse, "To catch a thief-you need at least 8 frames per second: the impact of frame rates on user performance in a CCTV detection task," in *Proc. 16th ACM Int. Conf. Multimedia*, pp. 941-944, Oct. 2008.
- [2] M. Shah and R. Garrad, "The development of large scale network-based multiuse CCTV surveillance systems," *IEE Int. Symp. on Imaging for Crime Detection and Prevention*, pp. 15-22, Jun. 2005
- [3] S. Yang and J. Park, "An efficient access control mechanism for secure surveillance systems," *J. KICS*, vol. 39B no. 4, pp. 228-233, Apr. 2014.

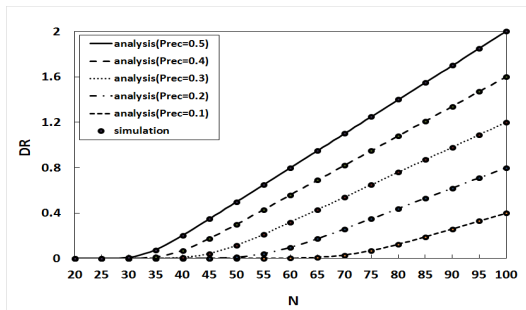


그림 2. $C=100$, $R_{rec}=5$, $R_{sil}=1$ 일 때 $P_{rec}=0.1, 0.2, 0.3, 0.4, 0.5$ 로 변환에 따른 N 과 DR 의 관계
Fig. 2. Relation between N and DR when $C=100$, $R_{rec}=5$, $R_{sil}=1$ and $P_{rec}=0.1, 0.2, 0.3, 0.4, 0.5$

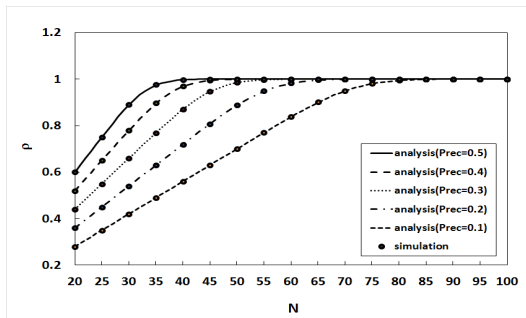


그림 3. $C=100$, $R_{rec}=5$, $R_{sil}=1$ 일 때 $P_{rec}=0.1, 0.2, 0.3, 0.4, 0.5$ 로 변환에 따른 N 과 ρ 의 관계
Fig. 3. Relation between N and ρ when $C=100$, $R_{rec}=5$, $R_{sil}=1$ and $P_{rec}=0.1, 0.2, 0.3, 0.4, 0.5$