

불법 스트리밍 영상 식별을 위한 유효 영역 및 프레임 기반 입력 영상 정규화 방법

박병찬*, 김영모*, 장의진**†

An Input Video Normalization Method Based on Valid Regions and Frames for Illegal Streaming Video Identification

Byeongchana Park*, Youngmo Kim*, Uijin Jang**†

요약

불법 스트리밍 영상은 재인코딩, 해상도 변환, 화면비 변경, 밝기 변화, 회전, 반전 등의 변형이 발생하며, 화면 녹화나 재송출 과정에서 검은 여백, 자막, 로고, 광고 영역과 같은 비콘텐츠 요소가 포함될 수 있다. 이러한 요소는 동일 콘텐츠에서 추출되는 특징 벡터의 일관성을 저하시켜 영상 식별 성능을 떨어뜨린다.

본 논문에서는 불법 스트리밍 영상 식별을 위한 유효 영역 및 프레임 기반 입력 영상 정규화 방법을 제안한다. 제안 방법은 유효 영역 검출, 종횡비 보존 해상도 정규화, 밝기 및 구조 정보 기반 프레임 정제, 회전 후보 기반 방향 보정, 유효 프레임 선별로 구성된다. 이를 통해 입력 영상의 불필요한 영역과 불안정한 프레임을 줄이고, 후속 특징 추출 과정에서 보다 일관된 특징 표현이 생성되도록 한다. 다양한 변형 조건이 적용된 질의 영상을 대상으로 실험한 결과, 제안 방법은 정규화 미적용 조건보다 영상 식별 인식률을 향상시켰으며, 대표 특징 벡터 구성을 통해 불필요한 특징 수를 줄이면서도 안정적인 식별 성능을 유지함을 확인하였다.

Abstract

Illegal streaming videos often undergo transformations such as re-encoding, resolution conversion, aspect ratio changes, brightness variation, rotation, and flipping. They may also include non-content regions such as black borders, subtitles, logos, and advertisements during screen recording or retransmission. These factors reduce the consistency of feature vectors extracted from the same content and degrade video identification performance. This paper proposes an input video normalization method based on valid regions and frames for illegal streaming video identification. The proposed method consists of valid region detection, aspect-ratio-preserving resolution normalization, frame refinement based on brightness and structural information, rotation candidate-based orientation correction, and valid frame selection. These steps reduce unnecessary regions and unstable frames, enabling more consistent feature representations in the subsequent feature extraction process.

Experimental results using query videos with various distortions show that the proposed method improves identification accuracy compared with the non-normalized condition. It also reduces unnecessary feature vectors while maintaining stable identification performance.

한글키워드 : 불법 스트리밍 영상, 영상 식별, 입력 영상 정규화, 유효 영역, 유효 프레임

keywords : illegal streaming video, video identification, input video normalization, valid region, valid frame

* 숭실대학교 컴퓨터학과

** 숭실대학교 스타르탄SW교육원

† 교신저자: 장의진(email: neon7624@ssu.ac.kr)

접수일자: 2026.05.30. 심사완료: 2026.06.13.

게재확정: 2026.06.20.

1. 서론

불법 스트리밍 영상은 정식 유통 경로를 거치지 않고 온라인 환경에서 무단으로 재송출되거나 공유되는 영상 콘텐츠로, 저작권 보호와 합법 콘텐츠 유통 질서를 저해하는 주요 요인이다[1], [3]. 이러한 불법 유통을 효과적으로 대응하기 위해서는 원본 콘텐츠와 불법 스트리밍 영상 간의 동일성 또는 유사성을 안정적으로 판단할 수 있는 영상 식별 기술이 필요하다[2], [4], [9].

그러나 불법 스트리밍 영상은 원본 영상과 동일한 내용을 포함하더라도 재인코딩, 해상도 변환, 화면비 변경, 밝기 변화, 회전, 반전, 프레임율 변화 등 다양한 변형을 포함할 수 있다[4], [5], [6]. 또한 화면 녹화나 재송출 과정에서 검은 여백, 자막, 로고, 광고 영역과 같은 비콘텐츠 요소가 추가되기도 한다[2], [6], [10]. 이러한 요소는 원본 콘텐츠의 시각적 구조를 변화시키고, 동일 콘텐츠에서 추출되는 특징 벡터의 일관성을 저하시켜 식별 성능 저하를 유발한다[5], [8], [10].

기존 영상 식별 연구는 주로 특징 추출 방법이나 검색 효율 향상에 초점을 두어 왔다[5], [7], [8]. 하지만 불법 스트리밍 환경에서는 특징 추출 이전 단계에서 입력 영상의 품질과 형태를 일정하게 정렬하는 과정이 중요하다[2], [4], [6]. 입력 영상에 포함된 불필요한 영역과 정보량이 부족한 프레임이 그대로 사용될 경우, 후속 특징 추출 과정에서 불안정한 특징 표현이 생성될 수 있기 때문이다[6], [8], [10].

본 논문에서는 불법 스트리밍 영상 식별을 위한 유효 영역 및 프레임 기반 입력 영상 정규화 방법을 제안한다. 제안 방법은 유효 영역 검출, 종횡비 보존 해상도 정규화, 밝기 및 구조 정보 기반 프레임 정제, 회전 후보 기반 방향 보정, 유효 프레임 선별로 구성된다.

본 논문의 구성은 다음과 같다. 제2장에서는

불법 스트리밍 영상의 왜곡 특성과 입력 정규화의 필요성을 설명한다. 제3장에서는 제안하는 입력 영상 정규화 방법을 기술한다. 제4장에서는 실험 환경과 결과를 제시하고, 제5장에서는 결론과 향후 연구 방향을 논의한다.

2. 불법 스트리밍 영상의 왜곡 특성과 입력 정규화 필요성

2.1 불법 스트리밍 영상의 입력 왜곡 특성

불법 스트리밍 영상은 원본 콘텐츠와 동일한 내용을 포함하더라도 수집, 변환, 재송출 과정에서 다양한 왜곡이 발생할 수 있다[1], [2], [4]. 대표적으로 재인코딩에 따른 화질 저하, 해상도 변환, 화면비 변경, 밝기 및 색상 변화, 회전, 반전, 프레임율 변화 등이 나타난다[5], [6], [10]. 또한 화면 녹화나 재송출 과정에서 검은 여백, 자막, 로고, 광고 영역과 같은 비콘텐츠 요소가 프레임에 포함될 수 있다[2], [4], [6].

이러한 왜곡은 원본 콘텐츠의 시각적 구조를 변화시키고, 동일 콘텐츠에서 추출되는 특징 벡터의 일관성을 저하시킨다[5], [8], [10]. 예를 들어 해상도와 화면비 변화는 객체의 상대적 크기와 공간 배치를 변화시키며, 회전과 반전은 프레임의 방향성을 다르게 만든다[5], [6]. 또한 검은 여백이나 광고 영역은 원본 콘텐츠와 무관한 특징을 생성하여 비교 과정에서 오차를 증가시킬 수 있다[2], [4]. 따라서 불법 스트리밍 영상 식별에서는 특징 추출 이전에 입력 영상의 형태와 품질을 일정하게 정렬하는 과정이 필요하다[2], [6], [8]. 이는 단순한 화질 보정이 아니라, 동일 콘텐츠에 대한 특징 표현의 변동성을 줄이기 위한 입력 정규화 과정이라 할 수 있다[5], [7], [10].

2.2 콘텐츠 기반 영상 식별과 입력 영상의 영향

콘텐츠 기반 영상 식별은 입력 영상에서 프레임 또는 구간 단위의 특징 벡터를 추출하고, 이를 원본 영상의 특징과 비교하여 동일 또는 유사 콘텐츠 여부를 판단하는 방식으로 수행된다[5], [8], [9]. 이 과정에서 입력 영상의 해상도, 방향, 밝기 분포, 유효 영역, 프레임 품질이 일정하지 않으면 동일 콘텐츠라 하더라도 서로 다른 특징 벡터가 생성될 수 있다[6], [8], [10].

기존 영상 식별 연구는 주로 강인한 특징 추출 방법이나 검색 효율 향상에 초점을 두어 왔다[5], [7], [8]. 그러나 불법 스트리밍 환경에서는 특징 추출 방법이 우수하더라도 입력 영상 자체에 비 콘텐츠 영역이나 심한 왜곡이 포함되어 있으면 식별 성능이 저하될 수 있다[2], [4], [6]. 따라서 입력 영상의 불안정 요소를 줄이는 정규화 과정은 후속 특징 추출의 안정성을 높이는 중요한 단계이다[5], [8], [10].

2.3 유효 영역 및 유효 프레임 기반 정규화 필요성

불법 스트리밍 영상에는 원본 콘텐츠와 무관한 영역이나 식별에 적합하지 않은 프레임이 포함될 수 있다[1], [2], [4]. 따라서 프레임 내 실제 콘텐츠가 포함된 유효 영역을 구분하고, 전체 프레임 중 식별에 유효한 프레임을 선별하는 과정이 필요하다[2], [6], [10].

유효 영역 검출은 검은 여백, 패딩, 광고 영역 등 비콘텐츠 요소의 영향을 줄이기 위한 과정이다[2], [4]. 이러한 영역을 그대로 사용할 경우 특징 벡터가 원본 콘텐츠 외부 요소의 영향을 받을 수 있으므로, 실제 영상 정보가 포함된 영역을 중심으로 정규화를 수행하는 것이 바람직하다[5], [6], [8].

또한 모든 프레임이 동일한 식별 정보를 제공하는 것은 아니다. 지나치게 어둡거나 밝은 프레임, 블러가 심한 프레임, 장면 전환 프레임, 중복

성이 높은 프레임은 특징 추출의 안정성과 처리 효율을 저하시킬 수 있다[4], [6], [9]. 따라서 밝기 정보, 구조 정보, 프레임 간 중복도 등을 고려하여 유효 프레임을 선별할 필요가 있다[5], [8], [10].

결과적으로 유효 영역 및 유효 프레임 기반 정규화는 불법 스트리밍 영상의 비정형 요소를 줄이고, 식별에 필요한 시각 정보를 안정적으로 유지하기 위한 과정이다[2], [5], [6]. 이를 통해 동일 콘텐츠에서 추출되는 특징 벡터의 일관성을 높이고, 영상 식별의 정확도와 검색 효율을 개선할 수 있다[8], [9], [10].

3. 유효 영역 및 프레임 기반 입력 영상 정규화 방법

3.1 개요

불법 스트리밍 영상은 재인코딩, 해상도 변환, 화면비 변경, 밝기 변화, 회전, 반전, 검은 여백, 자막, 로고, 광고 영역 등 다양한 비정형 요소를 포함할 수 있다. 이러한 요소는 동일 콘텐츠에서 추출되는 특징 벡터의 일관성을 저하시켜 영상 식별 성능을 떨어뜨린다. 따라서 본 논문에서는 특징 추출 이전 단계에서 입력 영상을 식별에 적합한 형태로 정규화하는 방법을 제안한다.

제안 방법은 입력 영상에서 프레임을 추출한 뒤, 유효 영역 검출, 해상도 정규화, 정보량 기반 프레임 정제, 회전 후보 기반 방향 보정, 유효 프레임 선별 및 대표 특징 벡터 구성의 순서로 수행된다. 전체 처리 절차는 그림 1과 같다.

입력 영상으로부터 추출된 전체 프레임 집합을 V 라 하면 식 (1)과 같이 나타낼 수 있다.

$$V = \{F_1, F_2, \dots, F_N\} \quad (1)$$

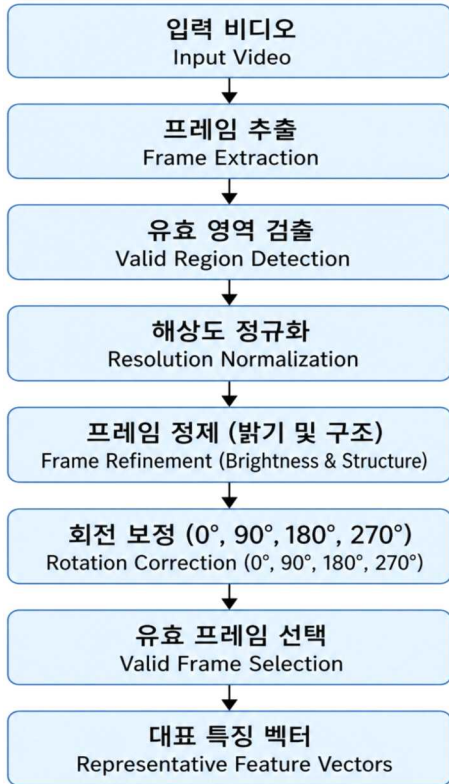


그림 1. 유효 영역 및 프레임 기반 입력 영상 정규화 절차

Fig. 1. Input video normalization procedure based on valid regions and frames

여기서 F_i 는 i 번째 프레임이고, N 은 전체 프레임 수를 의미한다. 즉 입력 영상은 시간 순서에 따라 배열된 프레임들의 집합으로 표현할 수 있다.

제안하는 전체 정규화 절차는 식 (2)와 같이 정의할 수 있다.

$$G = \Phi_{rep}(\Phi_{rot}(\Phi_q(\Phi_{res}(\Phi_{roi}(V)))))) \quad (2)$$

여기서 Φ_{roi} 는 유효 영역 검출, Φ_{res} 는 해상도 정규화, Φ_q 는 정보량 기반 프레임 정제, Φ_{rot} 는 회전 후보 기반 방향 보정, Φ_{rep} 는 대표 특징 벡

터 구성 과정을 의미한다. 최종 출력 G 는 비교·검색 단계에서 사용되는 대표 특징 벡터 집합이다.

3.2 유효 영역 검출

불법 스트리밍 영상에는 화면비 변환이나 화면 녹화 과정에서 검은 여백, 패딩, 광고 영역과 같은 비콘텐츠 영역이 포함될 수 있다. 이러한 영역은 원본 콘텐츠와 직접적인 관련이 없으므로 특징 추출 과정에서 불필요한 영향을 줄 수 있다. 따라서 본 논문에서는 프레임 내 실제 영상 정보가 포함된 유효 영역을 우선적으로 검출한다. 유효 영역 검출 과정은 그림 2와 같고, 처리 기준은 표 1과 같다.

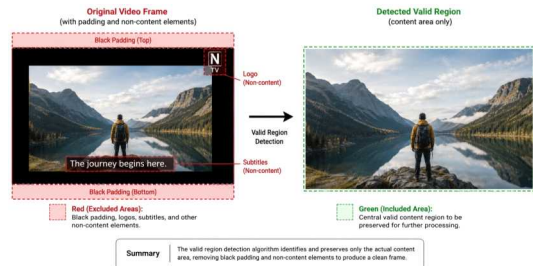


그림 2. 유효 영역 검출 예시
Fig. 2. Example of valid region detection

표 1. 유효 영역 검출 기준
Table. 1. Criteria for valid region detection

구분	조건	처리 방식
유효 영역	밝기값이 설정 범위 내에 존재하는 영역	유지
비콘텐츠 영역	지나치게 어둡거나 밝은 외곽 영역	제외
경계 상자	유효 영역 후보를 포함하는 최소 영역	추출 기준으로 사용

프레임 F_i 에서 유효 영역 후보 A_i 는 식 (3)과

같이 정의할 수 있다.

$$A_i = \{(x, y) | \tau_l < I_i(x, y) < \tau_u\} \quad (3)$$

여기서 $I_i(x, y)$ 는 프레임의 F_i 의 좌표 (x, y) 에서의 밝기값이며, τ_l 과 τ_u 는 각각 하한 및 상한 임계값이다. 지나치게 어둡거나 밝은 패딩 영역을 제외하고 실제 영상 정보가 존재할 가능성이 높은 영역을 찾기 위한 것이다.

검출된 유효 영역 후보를 기준으로 경계 상자 B_i 를 생성하고, 이를 이용하여 유효 영역 프레임 $F_i^{(c)}$ 를 식 (4)와 같이 생성한다

$$\begin{aligned} B_i &= BBox(A_i) \\ F_i^{(c)} &= Crop(F_i, B_i) \end{aligned} \quad (4)$$

여기서 $BBox(\cdot)$ 는 유효 영역 후보를 포함하는 경계 상자를 구하는 연산이며, $Crop(\cdot)$ 은 해당 영역을 기준으로 프레임을 잘라내는 연산이다.

3.3 해상도 정규화

불법 스트리밍 영상은 수집 환경이나 재송출 방식에 따라 서로 다른 해상도와 화면비를 가질 수 있다. 동일 콘텐츠라 하더라도 프레임 크기와 비율이 달라지면 객체의 상대적 크기와 공간 배치가 변화하여 특징 벡터의 일관성이 저하될 수 있다. 따라서 유효 영역이 검출된 프레임을 일정한 입력 크기로 정규화한다. 해상도 정규화 전후의 예시는 그림 3과 같고, 정규화 기준은 표 2와 같다.

유효 영역 프레임 $F_i^{(c)}$ 의 너비와 높이를 각각 w_i, h_i 라 하고, 목표 해상도를 $W \times H$ 라 할 때, 종횡비를 보존하기 위한 스케일 비율 s_i 는 식 (5)와 같이 정의한다.

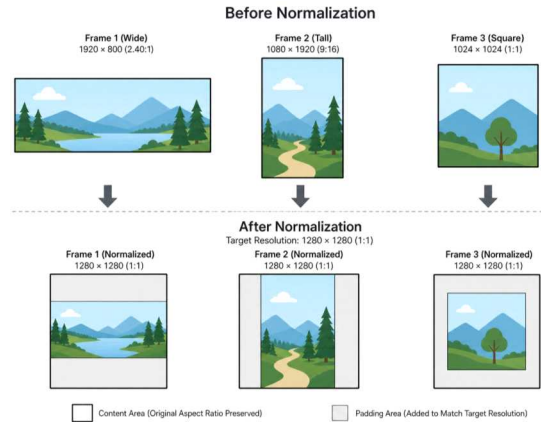


그림 3. 해상도 정규화 전후 비교
Fig. 3. Comparison before and after resolution normalization

표 2 해상도 정규화 기준
Table. 2 Criteria for resolution normalization

구분	정규화 전	정규화 후
프레임 크기	입력 영상별 상이	목표 해상도 $W \times H$
화면비	입력 영상별 상이	종횡비 보존
부족 영역	없음 또는 불규칙	패딩 적용

$$s_i = \min\left(\frac{W}{w_i}, \frac{H}{h_i}\right) \quad (5)$$

이를 이용하여 프레임을 확대 또는 축소한 뒤, 부족한 영역은 패딩으로 보완한다. 해상도 정규화 결과 $F_i^{(r)}$ 는 식 (6)과 같이 나타낼 수 있다.

$$F_i^{(r)} = P(S(F_i^{(c)}, s_i), W, H) \quad (6)$$

여기서 $S(\cdot)$ 는 스케일 조정 연산, $P(\cdot)$ 는 목표 해상도에 맞추기 위한 패딩 연산이다. 이 과정은 프레임의 종횡비 왜곡을 줄이면서 후속 특징 추출 과정에 일관된 입력 크기를 제공한다.

3.4 정보량 기반 프레임 정제

입력 영상의 모든 프레임이 식별에 유효한 정보를 제공하는 것은 아니다. 지나치게 어둡거나 밝은 프레임, 블러가 심한 프레임, 장면 전환 중간 프레임, 시각적 구조가 부족한 프레임은 특징 추출의 안정성을 저하시킬 수 있다. 따라서 본 논문에서는 밝기 정보와 구조 정보를 함께 고려하여 식별에 불리한 프레임을 제거한다. 정보량 기반 프레임 정제 과정은 그림 4와 같고, 프레임 유지 기준은 표 3과 같다.

표 3. 정보량 기반 프레임 유지 기준
Table 3. Criteria for frame refinement based on visual information

조건	의미	처리 방식
$Q_i \geq \tau_q$	식별에 필요한 정보량 충분	프레임 유지
$Q_i < \tau_q$	정보량 부족 또는 품질 저하	프레임 제거

먼저 정규화된 프레임 $F_i^{(r)}$ 의 밝기 히스토그램을 이용하여 확률분포 $p_i(k)$ 를 식 (7)과 같이 계산한다.

$$p_i(k) = \frac{h_i(k)}{\sum_{k=0}^{L-1} h_i(k)} \quad (7)$$

여기서 $h_i(k)$ 는 밝기값 k 에 해당하는 히스토그램 빈도이고, L 은 밝기 레벨 수이다. 이를 이용한 정규화 밝기 엔트로피 \hat{H}_i 는 식 (8)과 같다.

$$\hat{H}_i = \frac{-\sum_{k=0}^{L-1} p_i(k) \log p_i(k)}{\log L} \quad (8)$$

또한 프레임의 구조적 정보량을 반영하기 위해 에지 밀도 E_i 를 식 (9)와 같이 정의한다.

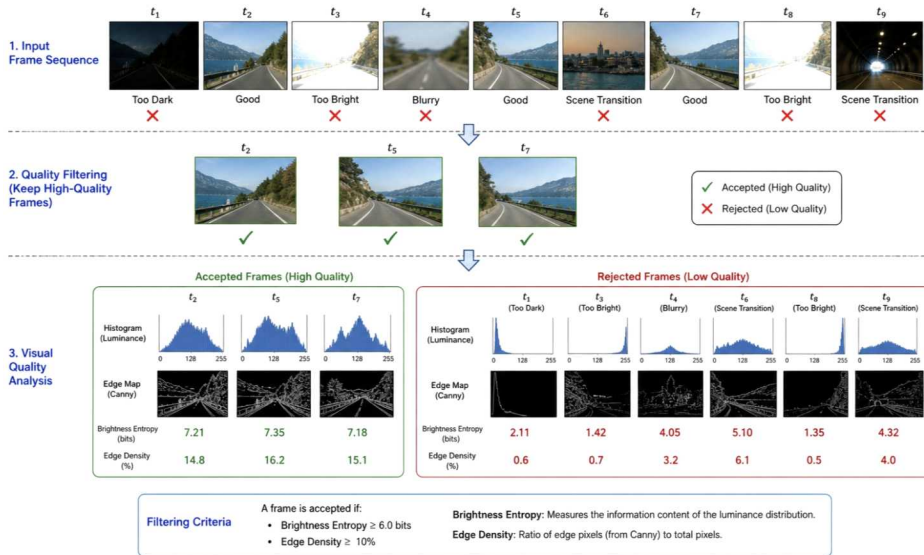


그림 4. 정보량 기반 프레임 정제 과정
Fig. 4. Frame refinement based on visual information

$$E_i = \frac{1}{WH} \sum_{x,y} 1(G_i(x,y) > \tau_e) \quad (9)$$

여기서 $G_i(x,y)$ 는 해당 위치의 에지 강도이며, τ_e 는 에지 판단 임계값이다. 최종 프레임 품질 점수 Q_i 는 식 (10)과 같이 계산한다.

$$Q_i = \alpha \hat{H}_i + \beta E_i \quad (10)$$

여기서 α 와 β 는 밝기 엔트로피와 에지 밀도의 반영 비율을 조절하는 가중치이다. 최종적으로 Q_i 가 임계값 τ_q 이상인 프레임만 유지하며 식 (11)과 같다. \emptyset

$$F_i^{(q)} = \begin{cases} F_i^{(r)}, & Q_i \geq \tau_q \\ \emptyset, & Q_i < \tau_q \end{cases} \quad (11)$$

이 과정은 단순히 프레임 수를 줄이기 위한 것이 아니라, 식별에 필요한 시각적 정보가 부족한 프레임을 제거하여 특징 추출의 안정성을 높이기 위한 것이다.

3.5 회전 후보 기반 방향 보정

불법 스트리밍 영상은 모바일 단말 환경이나 화면 녹화 과정에서 회전된 상태로 수집될 수 있다. 회전된 프레임은 동일 장면이라 하더라도 공간 배치와 방향성이 달라져 특징 벡터의 일관성을 저하시킬 수 있다. 따라서 본 논문에서는 0°, 90°, 180°, 270°의 회전 후보를 고려하여 방향 차이를 완화한다. 회전 후보 기반 방향 보정 과정은 그림 5와 같고, 회전 후보별 보정 기준은 표 4와 같다.

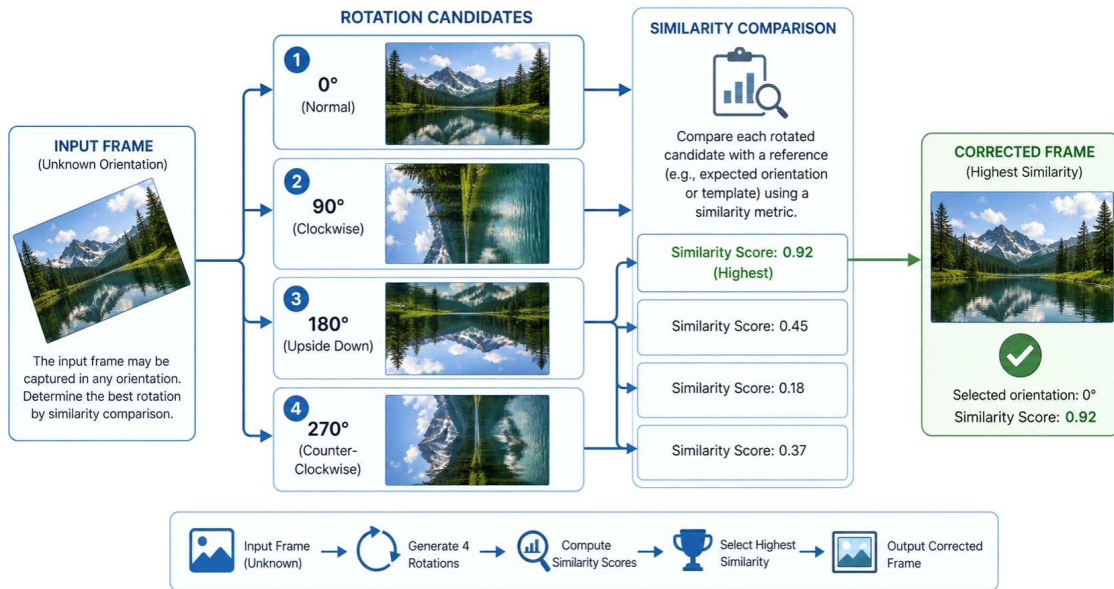


그림 5. 회전 후보 기반 방향 보정 과정
Fig. 5. Orientation correction based on rotation candidates

표 4. 회전 후보별 보정 기준
Table. 4. Correction criteria for rotation candidates

입력 상태	회전 후보	보정 방향
정상 방향	0°	보정 없음
시계 방향 회전	90°	-90°
반전	180°	-180°
반시계 방향 회전	270°	-270°

프레임 $F_i^{(q)}$ 에 대한 회전 후보 집합 R_i 는 식 (12)와 같이 정의한다.

$$R_i = \{T(F_i^{(q)}, \theta) | \theta \in \{0^\circ, 90^\circ, 180^\circ, 270^\circ\}\} \quad (12)$$

여기서 $T(\cdot)$ 는 회전 변환 연산이고, θ 는 회전 후보 각도이다. 각 회전 후보에서 특징 벡터를 추출하고, 기준 특징 또는 비교 대상 특징과의 유사도가 가장 높은 방향을 식 (13)에 의해 선택한다.

$$\theta_i^* = \arg \max_{\theta \in \{0^\circ, 90^\circ, 180^\circ, 270^\circ\}} Sim(z_{i,\theta}, z_{ref}) \quad (13)$$

여기서 $z_{i,\theta}$ 는 회전 후보 θ 에서 추출된 특징 벡터이며, $Sim(\cdot)$ 은 유사도 계산 함수이다. 최종 보정 프레임 $F_i^{(t)}$ 는 식 (14)와 같이 생성된다.

$$F_i^{(t)} = T(F_i^{(q)}, -\theta_i^*) \quad (14)$$

이와 같이 여러 회전 후보를 고려하면 단일 방향 추정 오류로 인한 성능 저하를 줄일 수 있으며, 회전 변형이 포함된 불법 스트리밍 영상에서도 보다 안정적인 특징 표현을 얻을 수 있다.

3.6 유효 프레임 선별 및 대표 특징 벡터 구성

정보량 기반 프레임 정제와 회전 후보 기반 방향 보정이 수행된 이후에도 연속 프레임 사이에는 중복성이 높게 나타날 수 있다. 이러한 중복 프레임을 모두 특징으로 저장하면 비교 연산량과 저장량이 증가한다. 따라서 본 논문에서는 유효 프레임으로부터 특징 벡터를 추출한 뒤, 코사인 유사도를 기준으로 대표 특징 벡터를 구성한다. 대표 특징 벡터 구성 과정은 그림 6과 같고, 특징 벡터 선별 기준은 표 5와 같다.

표 5 대표 특징 벡터 선별 기준
Table. 5. Criteria for selecting representative feature vectors

조건	의미	처리 방식
$\max \cos(z_i, g_j) < \tau_c$	기존 대표 벡터와 다름	대표 특징으로 추가
$\max \cos(z_i, g_j) \geq \tau_c$	기존 대표 벡터와 유사함	중복으로 판단하여 제외

회전 보정이 완료된 프레임 $F_i^{(t)}$ 에서 추출된 특징 벡터를 z_i 라 하면 식 (15)와 같이 나타낼 수 있다.

$$z_i = f(F_i^{(t)}) \quad (15)$$

여기서 $f(\cdot)$ 는 특징 추출 함수이다. 대표 특징 벡터 집합을 G 라 할 때, 새로 추출된 특징 벡터 z_i 가 기존 대표 특징 벡터들과 충분히 다름 경우에만 G 에 추가하며 식 (16)에 의해 수행된다.

$$\max_{g_j \in G} \cos(z_i, g_j) < \tau_c \quad (16)$$

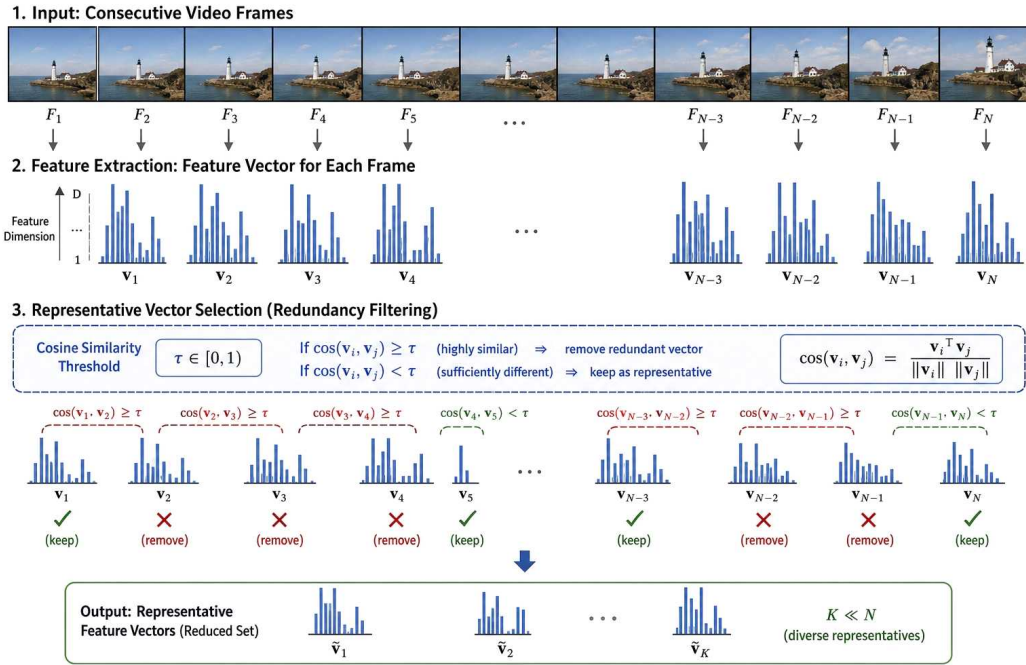


그림 6. 대표 특징 벡터 구성 과정
Fig. 6. Construction of representative feature vectors

조건을 만족하면 식 (17)과 같이 대표 특징 벡터 집합을 갱신한다.

$$G = G \cup \{z_j\} \quad (17)$$

여기서 $\cos(\cdot)$ 는 코사인 유사도이고, τ_c 는 중복 판단을 위한 유사도 임계값이다. 이 과정을 통해 연속 프레임에서 발생하는 중복 특징을 줄이고, 식별에 필요한 대표 특징만을 유지할 수 있다.

최종적으로 전체 프레임 수 대비 대표 특징 벡터의 압축률 CR 은 식 (18)과 같이 정의한다.

$$CR = \left(1 - \frac{|G|}{N}\right) \times 100 \quad (18)$$

여기서 $|G|$ 는 최종 대표 특징 벡터 수이고, N 은 입력 프레임 수이다. CR 은 대표 특징 벡터 구성을 통해 제거된 중복 특징의 비율을 백분율로 나타낸다.

4. 실험 및 결과

4.1 실험 환경

본 논문에서는 제안한 유효 영역 및 프레임 기반 입력 영상 정규화 방법의 효과를 검증하기 위해 실험을 수행하였다. 실험은 정규화 적용 전후의 식별 성능 비교, 대표 특징 벡터 구성을 위한 코사인 유사도 임계값 비교, 그리고 최종 성능 분석으로 구성하였다. 전체 실험 구성은 표 6과 같다.

4.2 실험 데이터 및 변형 조건

실험 데이터는 원본 영상과 다양한 변형이 적용된 질의 영상으로 구성하였다. 질의 영상에는 불법 스트리밍 환경에서 발생할 수 있는 해상도 변환, 화면비 변환, 회전, 반전, 밝기 변화, 색상

변화, 코덱 변환, 프레임을 변화 등을 적용하였다.

표 6. 실험 구성
Table 6. Experimental configuration

구분	실험 내용	목적
실험 1	정규화 적용 전후 성능 비교	입력 영상 정규화의 전체 효과 확인
실험 2	코사인 유사도 임계값 비교	대표 특징 벡터 구성 기준 확인
실험 3	최종 성능 분석	식별 성능과 특징 수 절감 효과 확인

실험 데이터 및 변형 조건은 표 7과 같으며, 원본 영상 데이터셋은 그림 7과 같다.

표 7. 실험 데이터 및 변형 조건
Table 7. Dataset and transformation conditions

구분	변형 유형
전체 프레임 수	18,000개
평가 단위	질의 영상 단위 및 프레임 특징 단위
비교 방식	코사인 유사도 기반 비교
코덱 변환	WMV, Xvid
화면비 변환	16:9 → 4:3, 4:3 → 16:9
해상도 변환	320×240, 640×480, 800×480
프레임을 변환	20fps
회전 변형	90도, 180도, 270도
반전 변형	수평 반전, 수직 반전
밝기 변형	±9%, ±18%
대조 변형	흑백 변환, 대조 62%, 137%

4.3 평가 방법

제안 방법의 성능은 인식률, 대표 특징 벡터 수, 압축률을 기준으로 평가하였다. 인식률은 전체 질의 영상 중 동일 원본 콘텐츠를 올바르게 식별한 질의 영상의 비율로 정의하며, 식 (19)와 같이 계산한다.

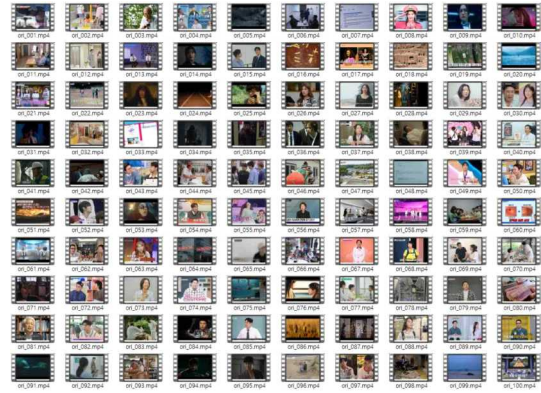


그림 7. 원본 영상 데이터셋
Fig. 7. Original video dataset

$$Accuracy = \frac{N_{correct}}{N_{query}} \times 100 \quad (19)$$

여기서 $N_{correct}$ 는 올바르게 식별된 질의 영상 수이고, N_{query} 는 전체 질의 영상 수이다.

대표 특징 벡터 구성에 따른 압축률은 전체 프레임 수 대비 최종 대표 특징 벡터 수의 감소 비율로 정의하며, 식 (20)과 같이 계산한다.

$$CR = (1 - \frac{N_{feature}}{N_{frame}}) \times 100 \quad (20)$$

여기서 $N_{feature}$ 는 대표 특징 벡터 수이고, N_{frame} 은 전체 프레임 수이다.

질의 영상과 원본 영상 간의 특징 비교에는 코사인 유사도를 사용하였다. 두 특징 벡터 z_i 와 z_j 의 코사인 유사도는 식 (21)과 같이 계산한다.

$$\cos(z_i, z_j) = \frac{z_i \cdot z_j}{\|z_i\| \|z_j\|} \quad (21)$$

코사인 유사도 값이 높을수록 두 특징 벡터가 유사함을 의미하며, 본 실험에서는 이를 기반으로

로 질의 영상과 원본 영상 간의 동일성 여부를 판단하였다.

4.4 정규화 적용 전후 성능 비교

제안한 입력 영상 정규화 방법의 효과를 확인하기 위해 정규화 미적용 조건과 정규화 적용 조건에서 식별 성능을 비교하였다. 정규화 적용 전후의 인식률 비교 결과는 그림 8과 같고, 세부 결과는 표 8과 같다.

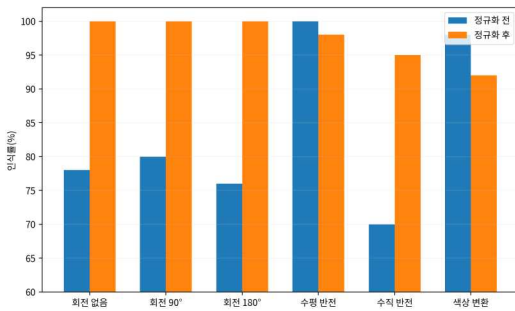


그림 8. 정규화 적용 전후 인식률 비교
Fig. 8. Comparison of identification accuracy before and after normalization

표 8. 정규화 적용 전후 성능 비교
Table 8. Performance comparison before and after normalization

조건	인식률 (%)	특징수 (개)	압축률 (%)
정규화 미적용	88.12	18,000	-
제안 방법 적용	98.93	8,224	54.31

실험 결과, 제안 방법을 적용한 경우 정규화 미적용 조건보다 인식률이 향상되는 것을 확인하였다. 이는 유효 영역 검출, 해상도 정규화, 정보량 기반 프레임 정제, 회전 후보 기반 방향 보정이 입력 영상의 비정형 요소를 줄이고, 후속 특징 추출 과정에서 보다 일관된 특징 벡터가 생성

되도록 기여했기 때문으로 해석할 수 있다.

또한 제안 방법은 전체 18,000개 프레임 중 8,224개의 대표 특징 벡터만을 사용하였으며, 이에 따라 54.31%의 특징 수 절감 효과를 보였다. 이는 제안 방법이 식별 성능 향상뿐만 아니라 비교 대상 특징 수를 줄여 처리 효율 개선에도 기여할 수 있음을 의미한다.

4.5 코사인 유사도 임계값 비교

대표 특징 벡터 구성 과정에서는 코사인 유사도 임계값에 따라 유지되는 특징 수와 인식률이 달라질 수 있다. 이에 따라 본 실험에서는 임계값을 0.4, 0.6, 0.8로 설정하고 각 조건에서 인식률, 대표 특징 수, 압축률을 비교하였다. 임계값에 따른 대표 특징 벡터 비교 결과는 표 9와 같다.

표 9. 임계값에 따른 대표 특징 벡터 비교
Table 9. Comparison of representative feature vectors according to cosine similarity threshold

코사인 유사도 임계값	프레임 수(개)	인식률 (%)	특징 수(개)	압축률 (%)
0.4	18,000	94.11	4,033	77.59
0.6	18,000	98.93	8,224	54.31
0.8	18,000	98.67	10,515	41.58

실험 결과, 임계값 0.4에서는 대표 특징 수가 4,033개로 가장 적고 압축률은 77.59%로 가장 높게 나타났다. 그러나 인식률은 94.11%로 다른 조건에 비해 낮았다. 이는 중복 제거 강도가 높아 식별에 필요한 특징 일부가 함께 제거되었기 때문으로 판단된다.

임계값 0.8에서는 인식률이 98.67%로 높게 유지되었으나, 대표 특징 수가 10,515개로 증가하여 압축률은 41.58%로 낮아졌다. 이는 많은 특징을

유지함으로써 식별 성능은 확보되지만, 저장량과 비교 연산량 측면에서는 효율이 낮아질 수 있음을 의미한다.

반면 임계값 0.6에서는 인식률이 98.93%로 가장 높았고, 대표 특징 수는 8,224개, 압축률은 54.31%로 나타났다. 따라서 인식률과 특징 수 절감 효과를 함께 고려할 때, 임계값 0.6이 가장 균형 잡힌 성능을 제공하는 것으로 판단된다.

5. 결론

본 논문에서는 불법 스트리밍 영상 식별 과정에서 입력 영상의 왜곡과 비콘텐츠 요소가 특징 벡터의 일관성을 저하시킨다는 문제를 해결하기 위해, 유효 영역 및 프레임 기반 입력 영상 정규화 방법을 제안하였다. 제안 방법은 유효 영역 검출, 중형비 보존 해상도 정규화, 정보량 기반 프레임 정제, 회전 후보 기반 방향 보정, 유효 프레임 선별 및 대표 특징 벡터 구성으로 이루어진다.

실험 결과, 제안 방법은 정규화 미적용 조건보다 영상 식별 성능을 향상시키는 것으로 나타났다. 특히 전체 18,000개 프레임 중 8,224개의 대표 특징 벡터만을 사용하여 54.31%의 압축률을 보이면서도 98.93%의 인식률을 유지하였다. 이는 제안 방법이 입력 영상의 비정형 요소를 줄이고, 후속 특징 추출 과정에서 보다 안정적인 특징 표현을 생성하는 데 효과적임을 보여준다.

또한 코사인 유사도 임계값 비교 결과, 임계값 0.6에서 인식률 98.93%, 압축률 54.31%로 정확도와 효율의 균형이 가장 우수하게 나타났다. 따라서 대표 특징 벡터 구성을 위한 기준으로 임계값 0.6이 적절함을 확인하였다.

향후 연구에서는 실제 불법 스트리밍 환경에서 수집된 다양한 콘텐츠와 대규모 질의 영상을 대상으로 제안 방법의 적용 가능성을 추가 검증

할 필요가 있다. 또한 자막, 워터마크, 광고 영역 등 원본 콘텐츠와 겹쳐 나타나는 비콘텐츠 요소를 보다 정밀하게 처리하는 방법에 대한 연구가 필요하다.

본 연구는 문화체육관광부 및 문화기술기획평가원의 2026년도 소프트웨어 저작권 사업으로 수행되었음(과제명 : OTT 콘텐츠 저작권 보호기술개발 및적용을 위한 저작권기술(+법) 융합인재양성, 과제번호 : RS-2023-00225267, 기여율: 100%)

참고 문헌

- [1] W.-S. Lee, S.-Y. Jang, I.-J. Yoo, B.-C. Park, S.-H. Shin, S.-Y. Kim, and Y.-M. Kim, "An End Credits-Based Illegal Streaming Video Recognition Method for OTT Content Identification," *Journal of Software Assessment and Valuation*, vol. 20, no. 4, pp. 225-232, Dec. 2024. DOI: 10.29056/jsav.2024.12.23
- [2] I. Yoo, S.-Y. Jang, B.-C. Park, S.-Y. Kim, and Y.-M. Kim, "An Illegal Streaming Video Identification Method Using Low-Frequency Components of Fast Fourier Transform," *Journal of Institute of Korean Electrical and Electronics Engineers*, vol. 28, no. 4, pp. 573-579, Dec. 2024. DOI: 10.7471/ikeee.2024.28.4.573
- [3] I. Yoo, J. Lee, B. Park, S. Kim, and Y. Kim, "A Method for Generating Signature Information to Determine Illegal Distribution of Cloud-Based Streaming Video," *Journal of Software Assessment and Valuation*, vol. 18, no. 2, pp. 239-246, 2022. DOI: 10.29056/jsav.2022.12.24
- [4] M.-S. Choi, "Efficient Video Matching Method for Illegal Video Detection," *Journal of Digital Convergence*, vol. 20,

- no. 1, pp. 179-188, 2022. DOI: <https://doi.org/10.14400/JDC.2022.20.1.179>
- [5] Z. Yu, X. Liang, L. Chen, X. Zhang, and Z. Tang, "Robust Video Hashing with DWT and Tensor SVD for Copy Detection," *Expert Systems with Applications*, vol. 304, Article 130669, 2026. DOI: 10.1016/j.eswa.2025.130669
- [6] K. Fojcik, P. Syga, and M. Klonowski, "Extremely Compact Video Representation for Efficient Near-Duplicates Detection," *Pattern Recognition*, vol. 158, Article 111016, 2025. DOI: 10.1016/j.patcog.2024.111016
- [7] Y. Zhou, X. Li, C. Xiong, H. Yao, and C. Qin, "A Survey of Perceptual Hashing for Multimedia," *ACM Transactions on Multimedia Computing, Communications, and Applications*, vol. 21, no. 7, Article 195, pp. 1-28, 2025. DOI: 10.1145/3727880
- [8] M. Yu, Z. Tang, H. Zhuang, X. Liang, Z. Li, and X. Zhang, "Robust Video Hashing with Non-Negative Tensor Factorization for Copy Detection," in *Proceedings of the 2024 International Conference on Multimedia Retrieval*, pp. 1094-1098, 2024. DOI: 10.1145/3652583.3657609
- [9] E. Pizzi et al., "The 2023 Video Similarity Dataset and Challenge," *Computer Vision and Image Understanding*, vol. 243, Article 103997, 2024. DOI: 10.1016/j.cviu.2024.103997
- [10] G. Zhao, C. Qin, X. Luo, X. Zhang, and C.-C. Chang, "Perceptual Robust Hashing for Video Copy Detection with Unsupervised Learning," in *Proceedings of the 2023 ACM Workshop on Information Hiding and Multimedia Security*, pp. 13-18, 2023. DOI: 0.1145/3577163.3595110

저자 소개



박병찬(Byeongchan Park)

2015.2 평생교육원 학점은행 졸업
2018.2 숭실대학교 컴퓨터학과 석사
2023.8 숭실대학교 컴퓨터학과 박사
2023.9-현재 : 숭실대학교 컴퓨터학부 초빙교수
<주관심분야> 저작권 보호 기술 및 이용 활성화



김영모(Youngmo Kim)

2003.2 대전대학교 컴퓨터공학과 졸업
2005.2 대전대학교 컴퓨터공학과 석사
2011.2 대전대학교 컴퓨터공학과 박사
2012-현재 : 숭실대학교 교수
<주관심분야> 저작권 보호 기술 및 이용 활성화



장의진(Uijin Jang)

1999.8 숭실대학교 컴퓨터학부 졸업
2002.8 숭실대학교 컴퓨터학과 석사
2011.2 숭실대학교 컴퓨터학과 박사
2018.9-현재 : 숭실대학교 스파르탄SW교육원 부교수
<주관심분야> 네트워크, 인공지능, 디지털 저작권 보호 기술 및 저작물 이용 활성화