

영상의 전역 특징과 이동객체의 지역 특징을 융합한 동영상 검색 디스크립터 설계

정병만¹ · 이규원^{2*}

A Descriptor Design for the Video Retrieval Combining the Global Feature of an Image and the Local of a Moving Object

Byung-man Jung¹ · Kyu-won Lee^{2*}

¹Information and Communications Engineering, Daejeon University, Daejeon 300-716, Korea

²Information and Communications Engineering, Daejeon University, Daejeon 300-716, Korea

요 약

실시간으로 입력되는 영상으로부터 이동객체의 움직임 특징을 이용하여 움직임 분석에 적합한 디스크립터를 제안한다. 배경과 이동객체를 분리하기 위하여 배경학습을 행한다. 연속적으로 추출된 이동객체의 1차 모멘트를 이용하여 각 객체별로 이동 궤적을 추출한다. 연결 리스트를 이용하여 객체별로 추출된 1차 모멘트를 관리한다. 디스크립터는 격자 형태로 미리 지정된 9개의 지점 근방에 포함되는 이동객체의 1차 모멘트 좌표와 객체가 화면에 출현하는 시작 프레임 번호, 화면에서 사라지는 마지막 프레임 번호로 구성된다. 제안하는 전역 및 지역 특징 융합 디스크립터에 의한 비디오 검색은 둘 중 하나의 특징을 사용하는 기존의 방법에 비하여 효과적임을 확인하였다.

ABSTRACT

A descriptor which is suitable for motion analysis by using the motion features of moving objects from the real time image sequence is proposed. To segment moving objects from the background, the background learning is performed. We extract motion trajectories of individual objects by using the sequence of the 1st order moment of moving objects. The center points of each object are managed by linked list. The descriptor includes the 1st order coordinates of moving object belong to neighbor of the pre-defined position in grid pattern, The start frame number which a moving object appeared in the scene and the end frame number which it disappeared. A video retrieval by the proposed descriptor combining global and local feature is more effective than conventional methods which adopt a single feature among global and local features.

키워드 : 기술어, 이동 객체, 이동 궤적, 영상 검색

Key word : Descriptor, moving object, motion trajectory, video retrieval

접수일자 : 2013. 09. 03 심사완료일자 : 2013. 09. 30 게재확정일자 : 2013. 10. 15

* **Corresponding Author** Kyu-Won Lee(E-mail: kwlee@dju.ac.kr, Tel:+82-42-280-2558)

Dept. of Information and Communications Engineering, Daejeon University, Daejeon 300-716, Korea

Open Access <http://dx.doi.org/10.6109/jkice.2014.18.1.142>

print ISSN: 2234-4772 online ISSN: 2288-4165

©This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License(<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.
Copyright © The Korea Institute of Information and Communication Engineering.

I. 서론

영상 검색 시스템은 사회 안전망 구축, 교통 시스템, 마케팅 등 여러 분야에서 사용되고 있으며 미리 정해진 검색 조건을 위한 시스템이 아닌 다양한 검색 조건이 가능한 지능형 검색 시스템으로 발전하고 있다.

효과적인 영상 검색 방법으로는 기존 방법인 텍스트 기반 방법 이외에 의미론 기반, 내용 기반이 있다. 의미론 기반 검색은 영상을 추상화된 특징으로 표현하는 방법으로 질의어의 구문이나 의미를 분석하여 검색을 하는 방법이다. 내용 기반 영상 검색은 영상 자체에서 특징을 자동으로 추출하고 검색하여 객관적이며 자동화된 영상검색이 가능하다[1]. 영상에 존재하는 다양한 특징에는 글로벌 특징과 로컬 특징으로 분류할 수 있다. 글로벌 특징을 이용하여 디스크립터를 설계할 경우 디스크립터의 처리량이 방대해지고 처리시간 또한 효과적이지 못하는 경우가 발생한다. 이러한 단점을 해결하기 위하여 로컬 영역을 이용한 디스크립터를 제안하였으나 영상 검색시에 질의에 응답이 오인되어 검색되는 점이 발생한다.

본 논문에서는 실시간으로 입력되는 영상으로부터 배경학습을 통하여 이동객체를 검출한 후, 객체의 중심좌표를 추출하여 로컬 특징자로 사용하고 전체 영역에서 9개의 체크 영역을 이용하여 글로벌 특징으로 사용하여 글로벌 특징의 정확성과 로컬 특징의 신속성을 이용한 강건한 디스크립터 설계 방법을 제안한다. 제2장에서는 글로벌 특징과 로컬 특징을 이용한 디스크립터 설계 방법을 설명하고, 제3장에서는 제안하는 글로벌과 로컬 특징자를 융합한 디스크립터 설계 방법을 설명하며 제4장에서는 제안하는 방법에 대한 실험 및 결과고찰을 기술하고 제5장에서는 결론 및 향후 연구를 제시한다.

II. 관련연구

일반적으로 이동객체에 기반을 두는 영상 검색 시스템은 객체 검출 단계, 객체 추적 단계, 객체 특징점 단계, 디스크립터 단계로 분류 할 수 있다. 영상 검색 기술은 객체의 특징 정보를 기반으로 하는 방법이다[2].

2.1. Color Correlogram

칼라 코렐로그램은 글로벌 특징을 이용한 알고리즘이다. 영상전체에서 일정한 거리를 가지는 화소간에 나타나는 칼라의 분포를 확률로 나타낸다. 칼라 코렐로그램은 일정거리에 있는 화소의 칼라 변화 정보를 포함하기 때문에 칼라와 에지정보를 포함 할 수 있다. 따라서 우수한 영상검색 결과를 얻기 위해서는 다양한 거리에서 대해 칼라 코렐로그램을 얻어야 한다.[3] 이때 전체적인 영상에 대하여 칼라 코렐로그램을 적용시에 많은 계산 시간이 필요하다. 또한 계산된 양만큼의 데이터를 저장하기 때문 데이터 관리에서 효율이 떨어진다.

2.2. 로컬 디스크립터

로컬 디스크립터는 특징점 기반의 방법으로 영상 내에서 왜곡에 강인함이 중시되는 기술이다. 영상의 특징점 주변 화소 정보를 활용하여 로컬 패치를 추출하고 로컬 패치로부터 특징 정보를 담고 있는 KeyPoint를 추출한다. 추출된 KeyPoint를 사용하여 로컬 디스크립터를 추출한다.

특징점 기반 영상 매칭 기술의 성능은 근본적으로 영상 왜곡(회전, 확대, 축소, 밝기 변화, 노이즈 등)에 대하여 특징점과 로컬 디스크립터의 강인성에 의존된다고 할 수 있다. 추출된 로컬 디스크립터는 검색의 대상이 되는 영상에서 현 영상만의 KeyPoint를 붙여 색인자의 역할을 할 수 있다.[4] 따라서 로컬 디스크립터는 글로벌 특징보다 비교적 빠른 처리 속도와 디스크립터의 처리데이터가 적어 데이터 관리가 용이하다. 하지만 다른 사이즈의 영상이나 전혀 다른 영상이지만 비슷한 키포인트를 생성하여 검색에 오류가 발생 할 수 있다.

III. 제안하는 시스템

본 논문에서 제안하는 전체 시스템 구성도는 그림 1과 같다.

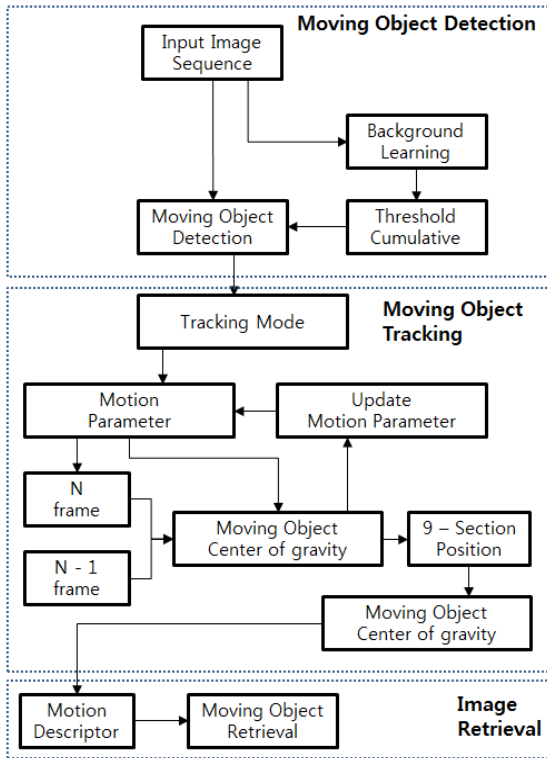


그림 1. 전체 시스템 구성도
Fig. 1 System Block Diagram

본 논문에서는 이동객체의 전역 및 지역 특징을 이용하여 디스크립터를 설계한다. 이동객체의 특징을 획득하기 위하여 배경학습모델을 기반으로 이동객체를 검출한다. 검출된 이동객체는 추적을 위해 객체의 무게중심 좌표를 현재 프레임과 이전 프레임의 연관성을 이용하여 하나의 단일 객체로 인지하고 각 이동객체 별로 하나의 연결리스트 구조를 생성하여 추적을 유지한다.

이동객체의 무게중심 좌표 값 생성 시에 임의의 지정한 9개의 영역에서만 생성되도록 한다. 이동객체의 추적이 끝나면 내용기반 영상 검색이 가능하도록 이동객체별 이동궤적 디스크립터를 생성하여 영상에서 추적된 객체의 검색을 가능하도록 한다. 이 장에서는 고정 카메라 환경에서 영상의 전역 특징과 이동객체의 지역 특징을 융합한 움직임 디스크립터 설계를 위해 제안한 알고리즘에 대하여 기술한다.

3.1. 배경학습 및 이동객체 검출

배경학습은 입력되는 영상을 프레임단위로 일정 시간동안 누적하고 누적되는 영상에서 한 프레임과 다른 프레임간의 차이를 절댓값으로 구하여 이를 누적하면서 배경학습 모델을 구성한다[5]. 절댓값을 구하는 식은 식 1과 같다.

$$h(x, y) = |f_t(x, y) - f_{t-1}(x, y)| \quad (1)$$

(1)에서 $h(x,y)$ 는 좌표 x, y 에 위치한 화소의 명도값 차이를 의미하고, $f_t(x,y)$ 는 현재 영상을 $f_{t-1}(x,y)$ 는 이전 영상을 의미한다. 누적된 영상들을 이용하여 배경에 대한 통계적 모델을 구한다. 통계적 모델은 각 픽셀에서 평균과 편차(평균 절댓값 차이)를 계산하고 이를 이용하여 입력되는 영상에서 배경과 전경을 분리한다. 따라서 이동객체 검출은 배경학습 모델에서 얻어진 임계값을 이용하여 객체가 출현 시 임계값이 존재하지 않는 픽셀로 간주 될 경우 그림 2와 같이 배경이 아닌 전경으로 인식되어 이동객체로 검출한다.

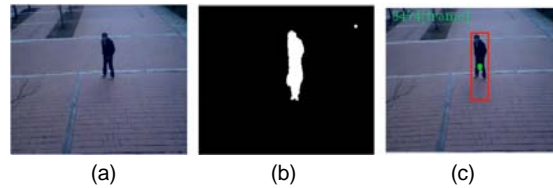


그림 2. 이동 객체 검출
Fig. 2 Moving Object Detection

그림 2의 (a)는 입력되는 원본영상이며, (b) 영상은 배경과 객체를 분리한 영상이며, (c)의 영상은 최종적으로 검출한 영상이다.

3.2. 이동객체 추적 및 디스크립터 생성

이동객체 추적은 검출된 이동객체를 기반으로 객체의 무게중심을 이용하여 추적한다. 무게중심은 (2)와 같이 검출된 객체의 x 값을 모두 더하고 y 값 또한 모두 더하여 검출된 픽셀의 수만큼 나누어 구한다.

$$G = \left(\frac{x_1 + x_2 + x_3}{n}, \frac{y_1 + y_2 + y_3}{n} \right) \quad (2)$$

추출된 무게중심은 좌표의 연관성을 이용하여 신규 객체 또는 이전 객체인지 유무를 판별하였다. 좌표의 연관성은 N프레임에서 검출된 이동객체의 무게중심 좌표와 N-1프레임에서 검출된 무게중심 좌표의 거리를 계산하여 근접으로 이동한 좌표를 같은 객체로 판별하였다. 좌표 거리 계산은 (3)과 같다.

$$\overline{XY} = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \quad (3)$$

좌표평면 위의 두 점 $X(x_1, y_1)$, $Y(x_2, y_2)$ 사이의 거리를 계산하여 기준을 정하고 짧은 거리로 이동한 객체를 동일 객체라 정의한다. 리스트에 객체의 무게중심 좌표를 저장하기 이전에 N프레임에서 추출한 좌표 값이 격자 형태로 미리 지정된 9개의 영역에 포함되는지 확인한다. 격자 형태로 미리 지정된 9개의 영역은 차후 영상검색에서 정확하고 효율적인 검색을 진행하기 위한 과정이다. 미리 지정된 9개 영역은 그림 3과 같다.



그림 3. 격자 형태로 미리 지정된 9개의 영역
Fig. 3 Nine regions defined

또한 이동객체 추적의 마지막은 객체가 출현하는 시점의 프레임 번호와 객체가 사라질 때까지의 프레임 구간을 정의한다. 프레임구간 정의는 영상 검색에 있어서 이동객체가 어떤 구간부터 어떤 구간까지 진행이 되었는지 정의 되어야하기 때문이다. 따라서 객체를 추적하면서 출현하는 시점의 프레임 번호는 유지하며 사라질 때 프레임 번호는 객체가 없어질 때 까지 계속 업데이트

트 하며 객체가 사라짐과 동시에 마지막 업데이트를 수행 하여 디스크립터를 설계한다. 설계된 디스크립터를 기반으로 영상검색을 수행한다. 영상검색 방법은 그림 4와 같다.

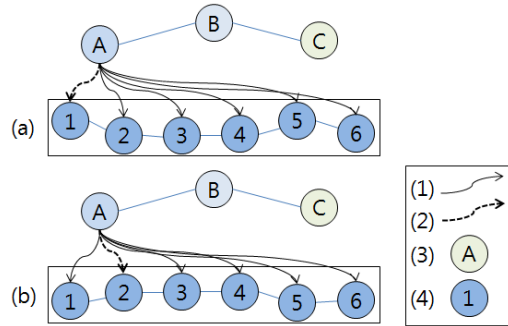


그림 4. 영상 검색 패턴
Fig. 4 Video search pattern
(1) 일반 거리
(2) 가장 근접 거리
(3) 질의 궤적
(4) Movement List 궤적

그림 4의 (a)는 (3)의 궤적을 갖는 질의가 입력되면 (4) 궤적을 이용하여 순차적으로 거리를 계산한다. (1)은 일반적인 거리를 나타내며 (2)는 질의 포인트와 가장 근접한 거리를 나타낸다. 이러한 거리비교 방법은 (3)의 B와 C도 같은 방법으로 비교하여 질의 궤적과 가장 유사한 궤적을 검출하여 출력한다. 또한 서로 다른 궤적을 가지고 있는 (a)와 (b)에 대하여 검색을 진행한 결과 가장 유사한 궤적을 선별하여 출력하고 우선순위에 따라 나머지 객체 궤적도 출력하였다.

IV. 실험 및 결과 고찰

실험에 사용된 영상은 실제 도로에서 디지털 카메라로 동영상 촬영을 한 실영상이다. 전체 프레임 수는 31,710프레임이며, 이는 취득된 영상에서 실험 시나리오에 해당하는 모든 구간의 프레임이다. 전체적인 실험은 크게 비슷한 방향으로의 궤적을 갖는 영상, 일정한 속도를 유지하는 영상, 속도의 변화폭이 존재 하는 영상으로 나뉘어 진행하며, 객체 검출, 객체 추적, 객체 히스토리 디스크립터 설계, 디스크립터를 이용한 객체 검

색으로 진행하였다. 사용된 이미지 해상도는 320x240이다. 그림 5는 주간 실외환경에 임의의 객체를 이동시켜 실험 동영상으로 사용한다.



그림 5. 실험영상
Fig. 5 Test Video

실험방법은 객체를 검출하고 추적하여 이동객체의 히스토리 정보를 저장한다. 표 1은 저장되는 히스토리 정보 테이블이다.

표 1. 이동객체 궤적 디스크립터
Table. 1 Moving object trajectory descriptor

Components	Value	Data Type
TO_ID	Object's ID	Int
Appearance Frame	Frame number	Int
Disappear Frame	Frame number	Int
Appearance Position	(x, y)	Point
Disappear Position	(x, y)	Point
Movement List	(x, y)	List
Object List	Movement List	List

실험 영상은 총 10개의 영상으로 실험하였으며, 실험 방법은 각 영상에 대해 검색 하고자 하는 궤적 모양의 좌표를 질의로 하여 가장 유사한 궤적을 가지는 객체를 Rank1부터 Rank3까지 응답하였다. Rank3까지 응답하는 이유는 실험 결과에 따른 성능평가에서 올바른 성능을 평가하기 위하여 제안 하였다. 따라서 가장 오차범위가 크게 나타날 수 있는 궤적을 시나리오-1, 시나리오-2로 분류하여 실험 하였다. 시나리오-1은 직진으로 출발하여 우회전 하는 궤적이다. 시나리오-2는 직진으로 출발하여 좌회전하는 궤적이다. 이외에 실험 영상은 좌측에서 우측, 우측에서 좌측으로 이동하는 객체를 담고 있지만 가장 정확하게 파악할 수 있는 실험영상을 시나리오로 작성한다. 시나리오-1에서 선택된 질의 방향에 따라 움직인 객체의 그림 4처럼 질의 패턴과 저장된 Moving object trajectory descriptor를 비교하여 질의 패턴과 저장된 이동객체의 움직임 패턴을 비교하여 가장 근사한 객체를 우선순위로 검색한다.

그림 6, 7은 시나리오-1, 2에 대한 질의 및 응답에 관한 결과 이다.



그림 6. 시나리오-1 검색결과 영상
Fig. 6 Scenario-1 results



그림 7. 시나리오-2 검색결과 영상
Fig. 7 Scenario-2 results

표 2. 제안하는 시스템의 성능 분석
Table. 2 Performance Analysis

실험 영상	전체 프레임 수	이동 객체 수	Recall (%)	Precision (%)
1	3670	7	100	40
2	3490	7	100	20
3	3470	6	100	60
4	4570	6	100	40
5	3538	7	100	60
6	3622	8	100	20
7	4076	10	100	60
8	1908	5	100	40
9	1719	5	100	60
10	1647	5	100	60
합계	31,710	66	평균(%)	
			100	46

표 2는 영상검색 시스템의 성능분석에서 사용되는 Recall과 Precision을 이용하여 제안하는 시스템의 성능 분석을 한 결과이다. 다음 (4,5)는 Recall 과 Precision을 구하는 식이다.

$$Recall = \frac{\sum_t S_t}{\sum_t O_t} \quad (4)$$

$$Precision = \frac{\sum_t L_t}{\sum_t P_t} \quad (5)$$

(4)의 O_t 는 실험영상에서의 질의 가능한 객체 수를 의미하며, S_t 는 질의를 통하여 응답된 리스트의 수를 의미한다.

(5)의 L_t 는 질의에 따른 응답된 리스트의 수를 의미하며, P_t 는 응답된 리스트 중 정확하게 응답된 객체를 의미한다.

제안하는 시스템과 비교를 수행하기 위하여 [6-7]에서 제안된 기법들과 비교한다.[6]은 컬러 레이아웃과 에지 히스토그램에서 얻는 특징자를 사용하여 검색하며, [7]은 JPEG2000에서 사용되는 웨이블릿 계수들의 멀티스케일 히스토그램을 기반으로 SIFT를 이용하여 비교한다. 표 3은 특성에 따른 비교표이다.

표 3. 모델 비교 분석
Table. 3 Model comparison

비교대상 비교항목	Ja-Hwng Su [6]	Cl. Moranda [7]	Proposed system
시각적 특징	○	○	○
궤적	×	×	방향, 거리
의미적 내용	동작	동작	동작
키 프레임	○	×	○
거리유사도	×	○	○

실험 결과 정확한 배경학습과 이동객체의 분할로 객체의 히스토리 정보를 이용하여 인식률 높은 영상검색 구현이 가능하다. 그러나 제안하는 시스템은 일반적으로 영상을 취득하는 카메라의 위치가 보행자 보다 높은 위치에 위치한다는 것에 착안하여 시스템을 구축하였고, 미리 제작된 시나리오를 이용하여 실험 환경 구축을 하였기 때문에 실험에 사용한 영상은 제안하는 시스템에서 적합하다고 할 수 있다. 하지만 많은 실험 영상을 통하여 실험을 하였고 320x240영상에서 24~28(fps)까지의 프레임속도를 유지한다는 점에서 제안하는 시스템이 실시간에 가깝게 처리 할 수 있음을 의미한다.

V. 결 론

본 논문에서 제안된 알고리즘은 이동객체 검출 및 추적 정보를 이용하여 히스토리 정보 디스크립터와 디스크립터 정보를 이용한 영상 검색이다. 9개의 체크 영역과 이동객체 중심좌표를 이용하여 디스크립터하고 검색환경에 맞도록 재구축하여 질의에 응답할 수 있도록 설계하였다. 제안된 알고리즘은 실험을 통하여 이동객체 움직임 정보 특징점 기반 디스크립터를 이용한 내용기반 영상 검색의 재현율, 정확성 측면에서 강건함을 증명 하였다. 향후 과제로는 본 알고리즘을 내용기반 영상 검색 시스템에서 객체의 다양한 정보를 디스크립터하고 제한적인 질의를 다양한 형태의 질의로 보완이 요구된다. 본 알고리즘은 지능형 영상 검색 시스템에 부합되도록 설계하였기 때문에 영상 데이터 검색, CCTV, 백화점, 대형 마켓 등 이동객체의 궤적 검색 및 분석을 필요로 하는 시스템에 적용이 기대된다.

REFERENCES

[1] Young-jun Song, "Video Content Search Technology Trends," *Korea Contents Association*, Vol. 7, No. 1, pp.

46-52, 2009.
 [2] N. Dimitrova and F. Golshani, "Motion recovery for video content classification," *ACM Trans Information Systems*, Vol. 13, No. 4, pp. 408-439, 1995.
 [3] Young-Eun An and Jong-An Park, "Color Correlogram using Combined RGB and HSV Color Spaces for Image Retrieval," *Korea Institute of Communications and Information Sciences*, Vol 32 No.5 pp.513-519, 2007.
 [4] Hyeok-June Jeon and Chi-Jung Hwang, "Image Matching (Correspondence Points) Based on Multi-resolution, Feature Point, Local Descriptor," *Korean Society for Internet Information Conference*, pp.5-848, 2010.
 [5] Myoung-kwan Oh, Kun-hee Han, Dong-jin Choi and Byung-min Jun "Moving Object Tracking using Differential Image," *Korea Contents Association Conference*, Vol. 2, No. 1, pp. 396-400, 2004.
 [6] Ja-Hwung Su, Yu-Ting Huang, Hsin-Ho Yeh and Vincent S. Tseng, "Effective content-based video retrieval using pattern-indexing and matching techniques," *Expert System with Applications*, Volume 37, Issue 7, pp. 5068-5085, 2010.
 [7] Cl. Moranda, J. Benois-Pineau, J. Ph. Domengera, J.Zepedab, E.Kijakc, and Ch. Guillemotb, "Scalable object-based video retrieval in HD video databases," *Signal Processing: Image Communication*, Volume 25, Issue 6, pp.450-465, 2010.



정병만(Byung-Man Jung)

2011년 대전대학교 정보통신공학과 졸업(학사)
 2013년 대전대학교 대학원 정보통신공학과(공학 석사)
 2013년 3월 ~ 현재 (주)한성시스코 연구소 연구원
 ※관심분야 : 영상처리, Robot Vision, 행동패턴



이규원(Kyu-Won Lee)

1986년 연세대학교 전자공학과 졸업(학사)
 1988년 연세대학교 대학원 전자공학과(공학 석사)
 1998년 연세대학교 대학원 전자공학과(공학 박사)
 1988년 2월 ~ 1989년8월 (주)LG산전연구소 연구원
 1989년 9월 ~ 2000년2월 한국전자통신연구원 선임연구원
 2003년 7월 ~ 2004년7월 Univ.of Massachusetts 방문 연구원
 2000년 ~ 현재 대전대학교 정보통신공학과 교수
 ※관심분야 : 영상처리, Motion Analysis, Visual Surveillance, Robot Vision, ITS