

# 반도체공정에서 영상처리를 사용한 바코드 문자 인식에 관한 연구

김경열\*, 서남원\*, 김수희\*  
\*호서대학교 컴퓨터공학과  
e-mail:kimkryul@nate.com

## A Study on the Barcode Character Recognition using Image Processing in Semiconductor Process

Kyung-Ryul Kim\*, Nam-Won Seo\*, Su-Hee Kim\*  
\*Dept of Computer Engineering, Ho-Seo University

### 요 약

반도체 공정에서 Track In/Out 판단을 위해서 바코드와 RFID(Radio-Frequency IDentification)가 주로 사용되고 있다. 현재 알려진 바코드 인식률은 약 80%로써 재고 유실이 여전히 존재하며 정확성과 신뢰성을 담보하지 못하고 있는 현실이다. 이 연구에서는 20% 정도의 인식되지 않은 부분을 보완하기 위해 바코드를 영상 처리하고 그 내부에 있는 문자들을 인식하는 프로그램을 개발한다. 해당 공정에서 바코드가 제대로 인식이 안 된 경우에는 인식된 문자들을 참고함으로써 공정의 전체 과정을 파악하고 트래킹할 수 있다. 바코드의 문자 인식을 생산 공정에 맞추어 실시간에 수향하여 필요시 활용함으로써 공정의 정확성과 생산성을 향상에 기여하고자 한다.

### 1. 서론

현재 우리나라는 반도체 메모리 분야에서는 세계 최고의 기술력과 경쟁력을 지니고 있다. 하지만, 반도체 공정에서 사용되고 있는 RFID 태그와 바코드를 인식하고 모니터링 하는 부분의 발전은 연구를 많이 하고 있지만, 연구에 비해 그에 따른 기술은 아직 미흡하다.

지금까지의 반도체 공정에서는 관리자가 직접 체크를 하거나 원거리 카메라를 이용하여 전체적인 생산 관리를 해왔다. 기록되지 않는 재고 유실을 추적할 수 없는 한계를 보완하기 위해 RFID 태그와 바코드를 사용하고 있지만, RFID 태그는 인식률이 97%로 높으나 가격이 비싸기 때문에 사용함에 있어 부담이 따른다. 반면에 바코드는 가격이 저렴하지만 인식률이 80%이다[1]. 이 연구에서는 가격이 저렴한 바코드를 사용하면서 작업자의 실수를 없애고 인식률을 100%로 만들기 위하여 바코드의 영상을 이용하여 바코드 문자를 인식하는 프로그램을 개발한다.

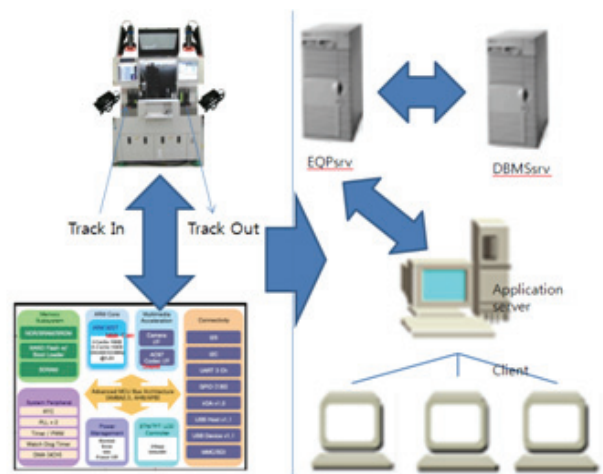
사실 문자인식에 대한 연구는 그 역사가 매우 깊으며 인쇄체인 경우는 인식률이 100%라는 것은 잘 알려진 사실이다. 그러나 이미 개발된 문자 인식 소프트웨어를 해당 공정에 사용하는 것이 용이하지는 않다. 또한 임베디드 PC에서 필요한 데이터의 처리에 대한 속도는 매우 중요하다.

이 논문에서는 문자 인식 분야의 기존 기술을 이용하여 바코드 문자 인식을 100%로 만들과 동시에 빠른 속도로 처리할 수 있도록 초점을 맞춘다. 개발하고자 하는 인식프

로그램을 실시간 반도체 공정을 모니터링하는 시스템의 맞춤형 문자인식 컴포넌트로 사용함으로써 공정의 정확성과 생산성 향상에 기여하고자 한다.

### 2. 실시간 반도체 공정 모니터링 시스템 구성도

반도체 공정에서 공정의 전반적인 과정을 실시간 모니터링하는 것은 매우 중요하다. 그렇지만 규모가 비교적 크지 않은 기업에서는 여전히 공정의 많은 데이터들이 수작업으로 처리하고 있는 현실이다.



[그림 1] 반도체 공정 모니터링 시스템 아키텍처

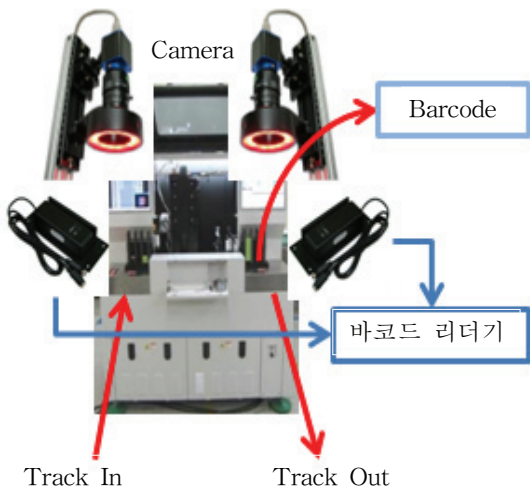
[그림 1]은 일반적인 반도체 공정에서 공정과정을 모니터

터링하는 시스템의 아키텍처이다. [그림 1]에서 보는 바와 같이 현재 처리되고 있는 공정에 대한 바코드 정보와 바코드 영상이 임베디드 컴퓨터에 저장되며, 동시에 대문 역할을 하는 서버에 실시간으로 전달된다. 대문 서버에서는 임베디드 컴퓨터에서 받은 데이터들을 데이터베이스에 저장한다. 어플리케이션 서버는 데이터베이스에 저장된 데이터를 실시간으로 웹상에서 확인하고 이상이 있는 경우 즉각적으로 대응할 수 있도록 한다.

**2-1. 바코드 인식과 영상처리 기술**

반도체 공정에서 제품이 들어가는 것을 Track In이라고 하고 나오는 것을 Track Out라고 한다. 아직도 일부 공정에서는 Track In과 Out을 하면서 작업자가 바코드를 리더기로 직접 인식하고, 제대로 인식을 못 하였을 경우에는 바코드 문자를 일일이 수작업으로 기록한다. 소요하는 시간적인 측면으로 보았을 때 매우 비효율적이고, 생산성이 떨어진다. 만약 바코드 인식이 제대로 되지 않을 경우를 대비하여 동일한 바코드를 영상으로 처리하여 문자 인식된 데이터를 저장하고, 필요시 활용한다면 작업자가 일일이 수작업으로 바코드 문자를 기록하는 수고를 덜어주며, 또한 생산성을 향상할 수 있다.

즉, 바코드 문자 인식을 소프트웨어적으로 처리함으로써 시간을 절약하게 되면 제품의 생산 효율을 향상시킬 수 있다. 그렇기 때문에 Track In/Out 시에 바코드 리더기와 카메라를 설치하고 바코드 리더기로 읽고 동시에 카메라로 이미지를 생성하여 바코드 내부에 있는 문자들을 인식하여 활용하는 것은 매우 필요하다.



[그림 2] 반도체 공정 설비 구성도

[그림 2]에서 보는 바와 같이 Track In 부분과 Track Out 부분에 카메라와 리더기를 설치하고 영상의 데이터와 바코드 데이터를 전송하고 처리하는 과정을 거칠 것이다.

**3. 바코드, NTSC 카메라**

우선적으로 이 연구에서는 바코드를 사용하여 리더기를 통한 인식에서 오류가 발생하는 상황을 보완하기 위해 NTSC 카메라를 사용한다. 인식 오류가 발생한 바코드를 확인하고 그 바코드에 해당하는 문자를 인식하여 저장하였다가 활용하면 재고 유실을 줄이고 생산성을 향상할 수 있다.

**3-1. 바코드**

바코드는 컴퓨터가 읽고 입력하기 쉬운 형태로 만들기 위하여, 문자를 흑과 백의 막대 기호와 조합한 코드를 말한다. 공장 자동화 분야에서는 가공 대상물에 바코드를 부착시켜 Lot 번호와 물품번호를 인식하여 작업상에 필요한 여러 사항을 파악하는데 도움을 준다[2].

현재 대부분의 반도체 공정에서 사용하는 바코드는 규격이 일정하고 문자 폰트와 크기가 모두 동일하다. 이 연구에서도 인식하고자 하는 바코드의 형태나 규격이 일정하다고 가정하고 인식 프로그램을 개발한다.

**3-2. NTSC 카메라**

카메라에는 기본적으로 NTSC 방식의 카메라와 PAL 방식의 카메라가 있다. 현재 NTSC 방식은 한국, 미국, 일본 등 44 개국이 사용하고 있는 비디오 출력 방식이며, PAL은 독일, 영국 등 100 여 개국이 사용하고 있다[3].

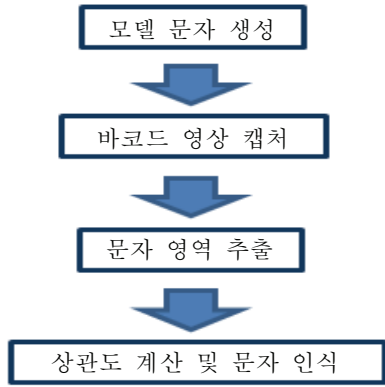
디지털 카메라의 기종에 따라서 NTSC와 PAL 방식을 두 가지 다 채용하거나 한 가지 방식을 채용하는 카메라가 있다. 이 두 가지 방식은 서로 호환이 되지 않으므로 제품 구입 시 두 가지 방식을 확인해야 한다.

NTSC 방식은 대형 카메라에 속해있는 것으로 크기가 크지만, PAL 방식에 비해서 노이즈가 적다는 장점이 있다. 자동 화이트 밸런스/DSP(디지털 신호 처리) 등이 되어서 문자를 인식할 때의 노이즈 감소에 도움이 되어 이 프로그램 개발을 위한 바코드 영상 촬영에 사용한다.

**4. 바코드 문자 인식기 개발**

일반적으로 바코드에는 숫자를 포함하는 문자가 들어가지만 한국에서 사용하는 바코드는 대부분은 숫자로 이루어져있다. 이 연구에서는 NTSC 카메라를 이용하여 바코드 영상을 캡처하고 문자 영역을 추출하여 그 내부의 문자들을 인식하는 프로그램을 개발하였다. 캡처한 영상은 8 비트 BMP 파일로 저장하였고, C언어의 콘솔 어플리케이션을 사용하여 문자를 인식하였다.

문자인식을 하는 방법은 여러 방법이 있지만 여기에서는 먼저 모델이 되는 문자들을 표준화하여 만든다. 인식하고자 하는 바코드를 대상으로 바코드 영역과 문자 영역을 분할한다. 추출한 문자 영역 전체를 픽셀단위로 스캔하면서 모델링한 문자들과 상관도를 계산한다. 상관도가 높은 경우는 해당위치에 모델 문자가 있음을 의미한다[4, 5].



[그림 3] 바코드 문자 인식 순서도

#### 4-1. 모델 문자 생성

가장 우선으로 바코드 문자 영역을 인식하기 위해 모델 문자를 생성한다. 생성된 모델 문자들과 바코드 문자 영역을 대상으로 상관도를 계산한다. 이 과정을 하기 위해서 문자 영역을 바코드 영역과 추출하여 모델이 되는 데이터를 [그림 4]와 같이 얻어낸다.

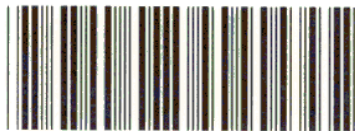


[그림 4] 모델이 되는 문자들

#### 4-2. 바코드 영상 캡처 및 영역 추출

##### ■ 바코드 영상 캡처

공정에서 사용하고 있는 바코드를 NTSC 카메라로 캡처하여 생성한다.



[그림 5] 캡처한 바코드

##### ■ 문자 영역 추출

캡처한 바코드를 대상으로 바코드 영역과 문자 영역을 구별하여 추출한다. 추출한 문자 영역 전체를 픽셀단위로 스캔하면서 모델링한 문자들과 상관도를 계산한다.



[그림 6] 추출한 문자 영역

#### 4-3. 상관도

문자 영역과 모델 문자들과의 상관도를 계산한다. 상관

도는 다음과 같이 계산할 수 있다[6].

$$r = \frac{S(xy)}{\sqrt{s(xx)s(yy)}}$$

여기에서

$x$  : 모델 문자의 픽셀 값

$y$  : 문자 영역에서 모델 문자와 대응하는 픽셀 값

$\bar{x}$  :  $x$ 값들의 평균,  $\bar{y}$  :  $y$ 값들의 평균

$$S(xx) = \sum_{i=1}^n (x_i - \bar{x})^2 : x \text{의 편차제곱의 합}$$

$$S(yy) = \sum_{i=1}^n (y_i - \bar{y})^2 : y \text{의 편차제곱의 합}$$

$$S(xy) = \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) : x \text{와 } y \text{의 편차 곱의 합}$$

#### 4-4. 상관도를 이용한 문자 인식

추출한 문자 영역을 대상으로 모델 문자를 0부터 9까지 하나씩 선택하여 픽셀 단위로 전체 영역을 이동하면서 각 지점에서 상관도를 계산한다.

계산한 지점에서 상관도 값이 크게 나온다는 것은 그 위치에 해당하는 문자와 유사한 패턴이 있다는 증거이다. 그러므로 상관도 값이 일정한 임계값을 초과하는 경우에 잠정적으로 그 위치에서 해당 문자가 있다고 판단한다. 예를 들어 0, 6, 9는 그 패턴이 상당히 유사하므로 모델 문자 0과 문자 영역에 있는 6 혹은 9와의 상관도는 매우 높을 수 있다. 동일한 위치에서 서로 다른 모델 문자와 상관도가 높게 나오는 경우에는 가장 상관도가 높게 나온 문자로 판별한다.



바코드 문자 인식 모듈



0 6 4 2 9 2 9

[그림 7] 바코드 문자 인식 모듈

[그림 8]은 개발한 프로그램을 이용하여 [그림 6]에 있는 문자 영역을 대상으로 문자들을 인식한 결과를 보여준다.

바코드에 있는 문자 영역의 0643551이 숫자 0 6 4 3 5 5 1로 판단한 것을 보여주고 있다. 적용한 상관도의 임계값은 0.8이다. [그림 8]의 첫 행은 픽셀 위치가 (79, 1)에서 모델 문자 0과의 상관도가 0.97이므로 그 위치에 숫자 0이 있다고 판단했음을 보여준다. 두 번째 행은 픽셀 위치가 (101, 1)에서 모델 문자 6과의 상관도가 0.92로 그 위치에 숫자 6이 있다고 판단했음을 보여준다. 다른 행도 동일한 방법으로 판단되었다.

x	y	corr	num
79	1	0.966768	0
101	1	0.922415	6
123	1	0.940548	4
146	1	0.867660	3
168	0	0.880960	5
191	0	0.900227	5
214	0	0.909723	1

Barcode Digits: 0 6 4 3 5 5 1  
계속하려면 아무 키나 누르십시오 . . .

[그림 8] 상관도 계산 화면 및 문자 출력 화면

### 5. 테스트 및 분석

4장에서 개발한 문자 인식 프로그램으로 20장의 바코드를 테스트 하였다. 바코드에 있는 이미지 문자들을 모두 정확하게 해당 숫자로 인식하였다.

동일한 이미지 문자 0이지만 바코드의 상태에 따라서 모델 문자 0과의 상관도는 상당히 차이는 경우도 있음을 관찰하였다. [표 1]에서 모델 문자 0과 이미지 문자 0의 상관도가 가장 높게는 0.97이고 가장 낮게는 0.8임을 볼 수 있다.

모델	0	1	2	3	4	5	6	7	8	9
상 관 도	0.97	0.91	0.97	0.87	0.94	0.88	0.92	0.99	0.92	0.84
	0.84	0.88	0.89	0.88	0.91	0.9	0.9	0.92	0.93	0.96
	0.87	0.95	0.89	0.92	0.93	0.95	0.89	0.89		0.97
	0.83	0.86	0.95	0.97	0.86	0.89	0.91	0.93		0.9
	0.85	0.91	0.94	0.94	0.9	0.88	0.91	0.94		0.93
	0.83	0.98	0.96	0.87	0.97	0.88	0.87	0.98		0.89
	0.81	0.87	0.87	0.85	0.89	0.92	0.8	0.93		0.93
	0.83	0.86	0.96	0.9	0.89	0.91	0.87	0.97		0.94
	0.84	0.84	0.93	0.9	0.89	0.92	0.87	0.95		0.86
	0.9	0.85	0.93	0.96	0.88		0.84			0.95
	0.93		0.94		0.92		0.93			0.91
	0.83		0.95		0.9		0.91			0.87
	0.81		0.94		0.98		0.91			0.9
	0.8		0.89		0.9		0.86			0.9
	0.87		0.93		0.87		0.91			0.92
	0.8		0.93		0.89		0.92			1
	0.88		0.93		0.93		0.94			
	0.83		0.89		0.9		0.82			
	0.85		0.92		0.9		0.86			
	0.81				0.93		0.95			
0.83				0.87		0.85				
						0.9				
평균	0.85	0.9	0.93	0.91	0.91	0.9	0.89	0.94	0.93	0.92

[표 1] 모델 문자와의 상관도 (Threshold : 0.8)

### 6. 결론 및 향후 과제

공정과정에서 바코드의 인식률은 80%로 알려져 있다. 이 연구에서는 20% 정도의 인식되지 않은 부분을 보완하기 위해 바코드를 영상 처리하고 그 내부에 있는 문자들을 인식하는 프로그램을 개발하였다.

개발한 문자 인식 프로그램으로 20장의 바코드를 테스트해 본 결과 정확도 100%로 이미지 문자들을 인식하였다. 시간상의 제약으로 오직 한 종류의 바코드 세트의 연구를 실행하였다. 앞으로는 다양한 글자체와 글자크기를 가지고 있는 바코드 인식 세트를 확보하여 테스트를 수행해서 그 결과를 바탕으로 취약한 부분들을 보완하고 확장하고자 한다.

개발한 프로그램은 궁극적으로 임베디드 컴퓨터에서 활용될 것이므로 각 세부 모듈별 수행 속도를 엄격하게 측정하여 처리 속도를 최대한 개선하고자 한다.

궁극적으로는 실시간 반도체 공정을 모니터링 하는 시스템의 맞춤형 문자인식 컴포넌트로 공정 작업 현장에서 사용될 수 있도록 신뢰성, 안정성, 신속성이 있는 문자 인식기로 개발하고자 한다.

### 참고문헌

- [1] [http://www.korpin.com/news\\_all/index.html?ln\\_mode=2\\_1&board=zboard37&mode=view&id=350&page\\_no=&ser\\_type=&search=&board\\_mode=](http://www.korpin.com/news_all/index.html?ln_mode=2_1&board=zboard37&mode=view&id=350&page_no=&ser_type=&search=&board_mode=)
- [2] <http://100.naver.com/100.nhn?docid=69140>
- [3] <http://www.dcinside.com/study/ntsc-pal.htm>
- [4] 강동중 / 하중은, Visual C++을 이용한 디지털 영상처리, 사이텍미디어, 2003. 03
- [5] Rafael C. Gonzalez / Richard E. Woods, Digital Image Processing Second Edition, Prentice Hall, 2003.09
- [6] 노형진, 통계적 품질관리, 한울출판사, 2008.06