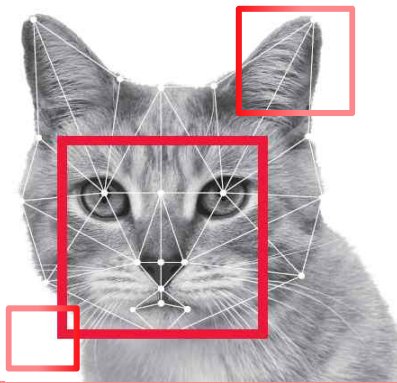


**한빛아카데미**  
Hanbit Academy, Inc.

**COMPUTER VISION** 컴퓨터 비전  
기본 개념부터 최신 모바일 응용 예까지



장 소개

## PREVIEW

- 오감 중에 시각은 가장 강력한 인지 기능
- 컴퓨터 비전은 컴퓨터를 이용하여 시각 기능을 갖는 기계 장치를 만드는 기술 분야



(a) 사진의 태동



(b) 초창기 컴퓨터 비전



(c) 현재 컴퓨터 비전  
그림 1-1 인공 시각을 만들기 위한 인류의 노력



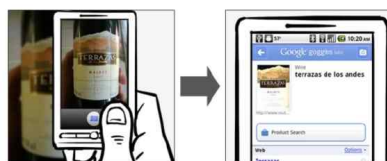
(d) 미래 컴퓨터 비전

## 각 절에서 다루는 내용

1. 왜 컴퓨터 비전인가?  
→ **최근 다양한 사례를 통해 컴퓨터 비전을 공부하는 이유를 설명한다.**
2. 컴퓨터 비전 문제는 어떻게 해결하나?  
→ **컴퓨터 비전 외에 다양한 연구들을 이차하는 접근 방법, 문제 해결에 사용하는 주요 도구를 살펴본다.**
3. 시스템 설계  
→ **컴퓨터 비전 시스템을 설계하고 구현할 때 가져야 하는 과제를 자세히 설명한다.**
4. 인접 학문  
→ **컴퓨터 비전과 인접한 학문인 영상 처리, 컴퓨터 그래픽스, 그리고 패턴 인식에 대해서 서로 어떻게 협력하고 어떻게 다른지 설명한다.**
5. 학습을 위한 자원  
→ **컴퓨터 비전을 공부하는데 도움이 될 책, 학술지, 학술대회, 웹 포털을 소개한다.**

### 1.1 왜 컴퓨터 비전인가?

#### ■ 생활 속의 응용 예



(a) 구글 앱



(b) Xbox 게임기

그림 1-2 외인 병을 인식하는 스마트폰 앱과 사용자의 동작을 인식하는 게임기

## 1.1 왜 컴퓨터 비전인가?

### ■ 팽창하는 응용

- 오락, 교통, 보안, 산업, 계산 사진학, 의료, 과학, 농업, 군사, 모바일 등



(a) 구글 자동차(자율 주행)



(b) 수술용 로봇 다빈치



(c) 화성 탐사선



(d) 딸기 따는 로봇



(e) 군사용 로봇 빅독



(f) 리프스냅

그림 1-3 컴퓨터 비전의 응용 사례

## 1.2 컴퓨터 비전 문제는 어떻게 해결하나?

121 과학적 접근과 공학적 접근

122 계층적 처리

123 문제 해결 도구

## 1.2 컴퓨터 비전 문제는 어떻게 해결하나?

- 사람의 시각은 빠르고 매우 강건
- 컴퓨터로 사람 수준을 달성하는 목표는 현재 기술로 불가능
  - 왜 어려운가?
  - 실용적인 시스템은 어떻게 만드나?

### 1.2.1 과학적 접근과 공학적 접근

- 목표 1 : 사람의 시각에 맞먹는 인공 시각을 만든다.



(a) 목표1 - 제약이 없는 상황

- 목표 2 : 한정된 범위에서 특정한 임무를 달성하는 인공 시각을 만든다.



(b) 목표2 - 제약이 있는 상황

그림 1-4 컴퓨터 비전의 목표

## 1.2.1 과학적 접근과 공학적 접근

### ■ 과학적 접근

- 목표 1을 달성하려는 노력
- 사람 시각의 원리를 밝혀낸 다음 컴퓨터로 모방
- 뇌 과학의 주요 관심사
- 지식 표현, 학습, 추론, 창작 같은 인공 지능이 필수

### ■ 어려운 이유

- 역 문제 (inverse problem)
- 불량 문제 (ill-posed problem)
- 다양한 변형 발생 (기하학적 변환, 광도 변환)

## 1.2.1 과학적 접근과 공학적 접근

### ■ 공학적 접근

- 목표 2를 달성하려는 노력
  - 특정 상황에서 특정 임무를 수행하는 실용 시스템 구축
- 성공적인 시스템
  - 많은 응용 현장에서 쓰고 있음
  - 컴퓨터 비전이 사람보다 뛰어난 경우 (예, 엔진 실린더 정밀 측정, 칩 검사 등)

### ■ 실용적인 성능 달성의 어려움

- 여전히 역 문제, 불량 문제, 다양한 변환 발생
- 영상은 숫자 배열 형태

126 134 125 122 127 127 120 130 138 136 138 140 130 127 127 130 133 126 138 133 127 139 134 130 125 121  
117 123 114 116 120 122 118 120 122 117 122 126 124 117 126 130 99 102 105 120 113 109 105 100 111  
109 110 105 102 112 123 130 135 147 171 151 164 153 174 157 139 124 107 90 82 87 86 82 80 89 89  
108 105 100 116 117 129 163 196 210 217 206 215 211 138 185 176 167 143 117 91 80 77 88 91 84 79  
107 103 102 100 146 170 200 150 170 195 138 141 135 123 118 125 134 143 127 121 84 85 86 82 81  
104 107 115 134 159 171 170 136 115 120 127 83 83 82 80 83 80 103 113 125 108 83 91 90 86 83  
107 120 127 160 150 125 138 150 167 174 115 99 94 103 98 89 87 91 104 103 99 97 95 84 96  
111 130 156 134 151 157 188 206 216 212 136 114 92 83 87 110 106 100 88 97 101 101 95 82 103 120  
130 146 164 165 185 213 219 210 212 196 190 108 123 127 127 123 111 121 134 145 132 130 147 159 160 171  
138 141 170 185 186 215 222 211 214 216 208 160 102 103 107 163 166 167 169 159 160 150 136 139  
142 163 171 190 190 204 218 213 207 214 218 213 204 195 182 189 183 179 161 159 163 171 160 169 187  
141 151 164 188 170 180 197 204 201 197 196 196 190 190 187 176 163 157 156 156 161 163 166 174 186 182  
144 151 160 185 183 176 187 182 196 189 184 179 177 174 165 156 151 140 163 177 182 186 200 203  
162 160 168 176 183 183 182 180 180 178 172 164 161 159 164 146 140 143 149 173 184 190 190 193 199 205  
159 148 178 170 202 206 197 184 187 176 170 167 172 179 160 176 168 203 215 212 206 204 202 204 205  
161 171 185 197 210 204 198 211 210 206 212 219 210 206 215 225 226 220 215 214 209 210 214 216 211 200

그림 1-5 숫자 배열 형태의 영상



### 1.2.2 계층적 처리



그림 1-6 컴퓨터 비전의 계층적 처리

#### ■ 전처리

- 주로 영상 처리

#### ■ 특징 추출

- 에지, 선분, 영역, 텍스처, 지역 특징 등을 검출하고 특징 벡터 추출

#### ■ 해석

- 응용에 따라 다양한 형태

### 1.2.3 문제 해결 도구

#### ■ 자료 구조와 알고리즘

- 배열, 트리, 그래프, 힙, 해싱, 탐색 트리 등
- 탐색 방법, 동적 프로그래밍, 한정 분기 등
- 고속 처리가 주된 관심

#### ■ 수학

- 선형 대수, 미적분학, 확률과 통계 등
- 최적화 문제 풀이

#### ■ 기계 학습

- 기계 학습을 활용하는 사례 급증
- 신경망, SVM, AdaBoost(에이더부스트), 임의 숲 등

## 1.3 시스템 설계

131 문제 이해

132 데이터베이스 수집

133 알고리즘 설계와 구현

134 성능 평가

## 1.3 시스템 설계

### ■ 순차 처리와 피드백



그림 1-8 컴퓨터 비전 시스템 설계 과정

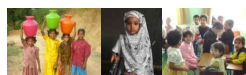
### 1.3.1 문제 이해

#### ■ 주어진 문제에 대한 직관적이고 철저한 이해 중요

- 합리적 제약 조건 수립 필요
- 예) 얼굴 인식기
  - 크기의 일정한 정면 얼굴 vs. 자연 영상 속의 얼굴
  - 전자는 제작 쉽지만 응용 분야 한정. 보안 장치에 활용 가능하지만 사진 분류 응용에 적용 불가능



VS.



### 1.3.2 데이터베이스 수집

#### ■ 데이터베이스

- 질적으로 양적으로 우수해야 고성능 시스템 제작 가능
- 데이터베이스=학습 집합+테스트 집합
- 수집 방법
  - 직접 수집 (많은 비용 부담해야 하지만 개발자에게 자산)
  - 인터넷에서 다운로드 (고품질의 데이터베이스 풍부함. 부록 B 참조)



(a) 물체 인식



(b) 차량 번호판 인식



(c) 영역 분할

그림 1-9 여러 가지 용도의 데이터베이스



### 1.3.3 알고리즘 설계와 구현

#### ■ 알고리즘 선택의 중요성과 어려움

- 새로운 알고리즘 개발 또는 기존 알고리즘 중에서 주어진 문제에 적합한 것 선택
- 예) 손 동작 인식기 제작
  - 에지, 영역, 지역 특징 중에 어떤 것을 사용할까?
  - 영역을 사용한다면, 어떤 영상 분할 알고리즘을 사용할까? 어떤 추적 알고리즘을 사용할까?

#### ■ 선택 방법

- 데이터베이스를 이용하여 성능 실험을 해 봄 (설계자의 경험과 직관이 중요)
- 성능 비교 분석 논문을 참조

### 1.3.3 알고리즘 설계와 구현

#### ■ 프로그래밍

- OpenCV (부록 A)
  - 국내 커뮤니티 <http://cafe.naver.com/opencv>
- Matlab이 제공하는 IPT (Image Processing Toolbox)
- 오픈 소스 (부록 B 참조)

### 1.3.4 성능 평가

#### ■ 인식 성능 측정

$$\begin{aligned} \text{정인식률} &= \frac{c}{N}, \text{기각률} = \frac{r}{N}, \text{오류율} = \frac{e}{N} \\ \text{이때 } c &= \text{맞는 샘플수}, r = \text{기각한 샘플수}, e = \text{틀린 샘플수} \quad (1.1) \\ (N &= c + r + e) \end{aligned}$$

- 부류가 심한 불균형일 때 부적절  
(예, 칩 검사에서 불량률이 0.1%라면, 임의의 짐작 분류기의 정인식률은 99.9%)

#### ■ 혼동 행렬

- 오류 경향을 세밀하게 분석하는데 사용

표 1-1 부류가 두 개인 경우의 혼동 행렬

참 부류 \ 분류 결과	$\omega_1$	$\omega_2$
$\omega_1$	$n_{11}(\text{TP})$	$n_{12}(\text{FN})$
$\omega_2$	$n_{21}(\text{FP})$	$n_{22}(\text{TN})$

### 1.3.4 성능 평가

#### ■ 참/거짓 긍정률, 참/거짓 부정률, 재현률과 정확률, $F$ 측정

$$\begin{aligned} \text{거짓 긍정률} \left( \text{FPR} = \frac{\text{FP}}{\text{FP} + \text{TN}} \right) &= \frac{n_{21}}{(n_{21} + n_{22})} \\ \text{거짓 부정률} \left( \text{FNR} = \frac{\text{FN}}{\text{TP} + \text{FN}} \right) &= \frac{n_{12}}{(n_{11} + n_{12})} \\ \text{참 긍정률} \left( \text{TPR} = \frac{\text{TP}}{\text{TP} + \text{FN}} \right) &= \frac{n_{11}}{(n_{11} + n_{12})} \\ \text{참 부정률} \left( \text{TNR} = \frac{\text{TN}}{\text{FP} + \text{TN}} \right) &= \frac{n_{22}}{(n_{21} + n_{22})} \end{aligned} \quad (1.2)$$

$$\begin{aligned} \text{정확률} &= \frac{n_{11}}{(n_{11} + n_{21})} \\ \text{재현율} &= \frac{n_{11}}{(n_{11} + n_{12})} \end{aligned} \quad (1.3)$$

$$\begin{aligned} F_\beta &= (1 + \beta^2) \frac{\text{정확률} \times \text{재현율}}{\beta^2 \times \text{정확률} + \text{재현율}} \\ F_1 &= \frac{2 \times \text{정확률} \times \text{재현율}}{\text{정확률} + \text{재현율}} \end{aligned} \quad (1.4)$$

표 1-1 부류가 두 개인 경우의 혼동 행렬

참 부류 \ 분류 결과	$\omega_1$	$\omega_2$
$\omega_1$	$n_{11}(\text{TP})$	$n_{12}(\text{FN})$
$\omega_2$	$n_{21}(\text{FP})$	$n_{22}(\text{TN})$

1.3.4 성능 평가

예제 1-1 얼굴 검출의 성능 측정

[그림 1-10]은 세 개의 영상을 가진 데이터베이스에서 얼굴을 검출한 결과를 보여준다. 앞에서 다룬 정확률, 재현율,  $F_1$  측정을 구해보자.



그림 1-10 얼굴 검출 성능

총 15개의 얼굴 중 12개를 옳게 검출했으므로 참 긍정  $n_{11}=12$ , 세 개의 얼굴을 못 찾았으므로 거짓 부정  $n_{12}=3$ , 그리고 얼굴 아닌 곳을 얼굴로 검출한 것이 두 개이므로 거짓 긍정  $n_{21}=2$ 이다. 따라서 정확률은  $12/14$ 이고 재현율은  $12/15$ 이다.  $F_1$  측정은  $24/29$ 이다.

1.4 인접 학문

■ 상호 협력이 강해지는 추세

- 예) 영화 제작, 계산 사진학 등

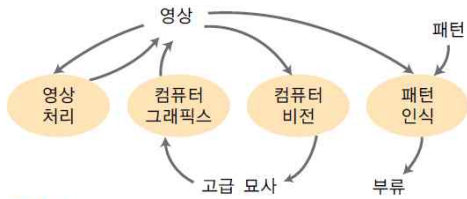


그림 1-11 인접 학문 간의 관계

## 1.5 학습을 위한 자원

### ■ 도서

- [Szeliski2011] Richard Szeliski, Computer Vision: Algorithms and Applications, Springer
  - 컴퓨터 비전의 최신 주제를 거의 빠짐없이 다룬다. 실용적인 응용을 군데군데 제시하여 이론과 응용을 균형 있게 설명한다.
  - 책의 원고 전체가 저자의 홈페이지에 공개되어 있다.
  - 하지만 알고리즘의 기초 원리부터 체계적으로 공부해야 하는 초보자를 위한 입문서로는 부족한 면이 있다.
- [Sonka2008] Milan Sonka, Vaclav Hlavac, and Roger Boyle, Image Processing, Analysis, and MachineVision, 3rd Edition, Thomson
  - 초보자용 입문서로 적당하다. Matlab 코드를 담은 쌍둥이 책인 [Svoboda2008]이있기 때문에, 컴퓨터 비전과 Matlab 프로그래밍 공부를 동시에 하여 일석이조의 효과를 거두고자 할 때 안성맞춤이다.
- [Shapiro2001] Linda G. Shapiro and George C. Stockman, Computer Vision, Prentice Hall
  - 초보자용 입문서 또는 대학 강좌의 교재로서 훌륭하다. 하지만 출판 이후 10년이 넘어 최신 주제와 알고리즘이 빠진 것이 흠이다.

## 1.5 학습을 위한 자원

### ■ 학술지

- **IEEE Transactions on Pattern Analysis and Machine Intelligence(PAMI)**
- **International Journal of Computer Vision(IJCV)**
- Image and Vision Computing
- Computer Vision and Image Understanding
- Foundations and Trends in Computer Graphics and Vision
- Pattern Recognition
- IEEE Transactions on Image Processing
- ACM Transactions on Graphics

## 1.5 학습을 위한 자원

### ■ 학술대회

- IEEE International Conference on Computer Vision(ICCV)
- IEEE International Conference on Computer Vision and Pattern Recognition(CVPR)
- European Conference on Computer Vision
- Asian Conference on Computer Vision
- British Machine Vision Conference
- International Conference on Pattern Recognition
- IEEE International Conference on Image Processing
- ACM SIGGRAPH

## 1.5 학습을 위한 자원

### ■ 웹 사이트(부록 B)

- CVonline(<http://homepages.inf.ed.ac.uk/rbf/CVonline/>)
  - 컴퓨터 비전의 주제를 계층적으로 나누고, 그들 주제에 대한 설명을 위키피디아를 이용하여 제공한다.
  - 데이터베이스, 문헌, 소프트웨어, 교육을 위한 최신 자료도 풍부하다.
- VisionBib.Com(<http://www.visionbib.com/bibliography/contents.html>)
  - 컴퓨터 비전과 관련된 논문을 망라하여 제공한다.
  - 주제어, 저자, 시기, 학술지, 학술대회에 따라 검색이 가능하다.
- Computer Vision Online(<http://www.computervisiononline.com/>)
  - 소프트웨어, 데이터베이스, 그리고 책에 대한 최신 정보를 제공한다.
  - 컴퓨터 비전 분야에서 주목할 만한 뉴스도 알려준다.