Study on mobile image search

Ioannis Kompatsiaris
CERTH-ITI

Spiros Nikolopoulos, CERTH-ITI, Stavri G Nikolov, IPTS, EC

NEM SUMMIT
Torino-Italy, 29th September 2011

Overview

- What is visual-based, mobile image search
- Moving mainstream and gaining momentum
- Architecture & Technologies
  - Client/server communication, client interfaces
  - Processing load, image recognition
- User needs and usage
- Business models
  - SaaS-based, contract-based, advertising-based
- Case study
- Service comparison and future trends

What is visual-based, mobile image search

source: Girod et al., "Mobile Visual Search", IEEE Signal processing magazine, 2011

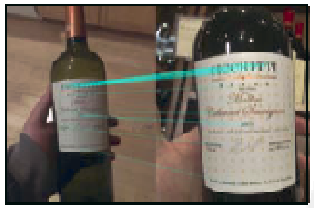
- Links the physical and digital worlds
- Makes the visual surroundings "**searchable**" and objects in visual proximity "**clickable**"

- Users turn their phone's camera towards the item of interest and further information (e.g. the **wine price**) or services (e.g. **buy the wine from the closest store**) are provided.

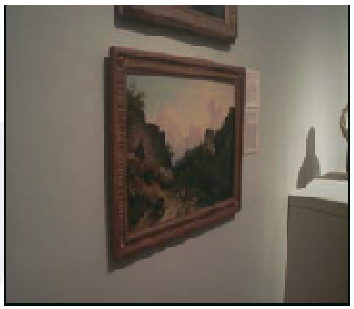
source: http://ideeinc.com/products/tineyemobile/

Mobile visual search applications

Ads/Catalogs

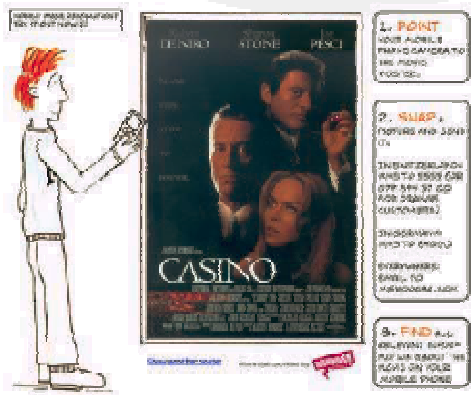


Wine Labels

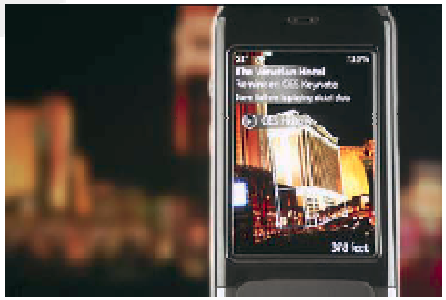


Museum Guide

Comparison Shopping



Movie Posters



Landmarks

Real Estate

Media covers
(CD, DVD, books)

source: Girod et al., “Mobile Visual Search”, IEEE Signal processing magazine, 2011

Moving mainstream and gaining momentum

Camera-enhanced devices are spreading at a very fast pace

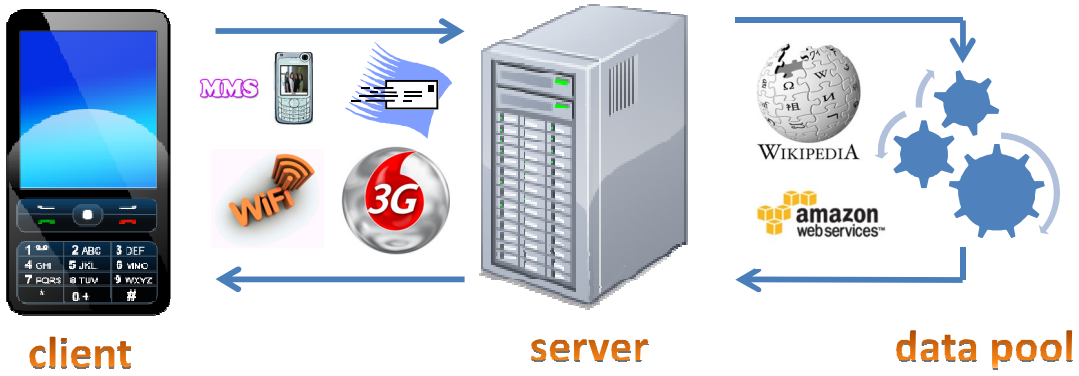
- In 2006 smartphones acounted only for **6.9%** of the total market, while in 2007 the market segment reached **10.6%.**
- The total sales of mobile devices reached **1,275 million** units in 2008 of which **15%** correspond to smartphones.
- In Europe **280 million** units were sold in 2008, of which **19,3%** where smartphones.

Users are embracing mobile image search

- Among **5013** internet users (end of 2010); **82%** notice mobile adds; **74%** make a purchase using their smartphones, **88%** that look for local information in their smartphones, take action within a day.
- In the end of April 2010 the mobile impressions account for **10,2%** of all paid search impressions (desktop+mobile)

eMarketer predicted that by 2011 the mobile image search market will account for around $715 million and we can reasonably expect that the visual-based image search segment will scale proportionally.

Architecture



- Client/server architecture
- Smartphones clients that capture the images and send the queries
- Server recognizes the image content, retrieves the relevant information from the data pool and sends back the response

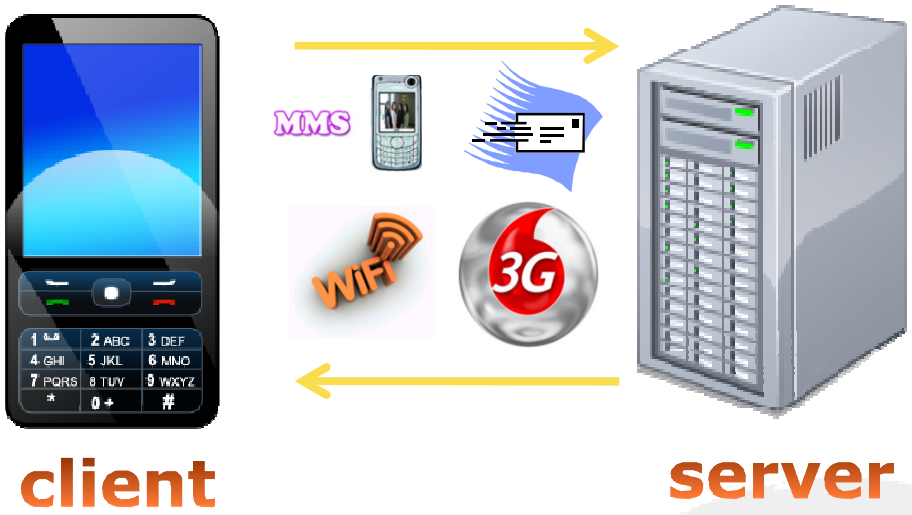
Important technological & design aspects

- Client/server communication
- Client interface
- Processing load
- Image content recognition

Technologies I

Client/server communication

- Using **MMS (multimedia messaging)** or **e-mail** services
  - Used primarily on older cell phones due to the low speed capabilities **GSM/GPRS** networks.
- **Wifi** or **3G networks**
  - Transparent to the user
  - Applicable in **Wifi** or **3G networks** and is adopted by the most recent image search applications.



Client interface

- **Menu mediated interaction**, switch between different menus
- **Snap-based interaction**, a single interface is used
- **Real-time interaction**, an augmented reality experience with meta-tags popping up as the camera turns towards the object of interest.

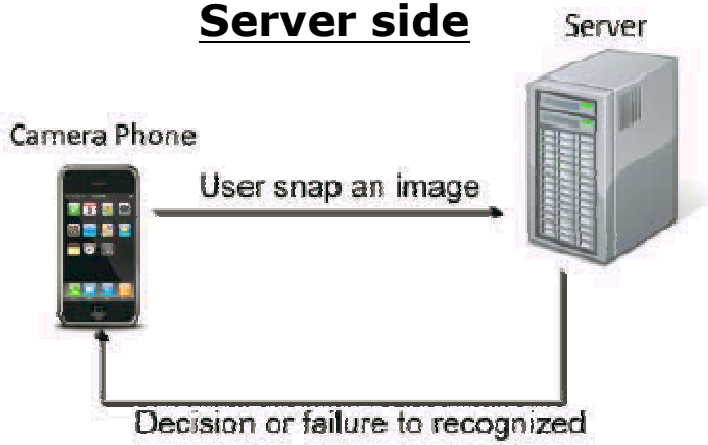
source:
http://www.extratech.gr/

Technologies III

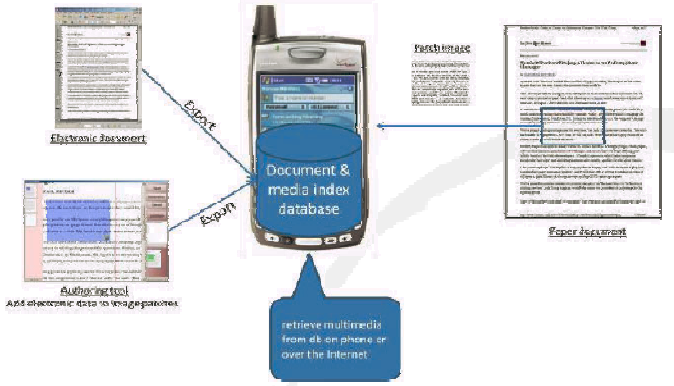
Processing load

- **Client-side**
  - the captured image is processed by the smartphone and only a representative set of features is transmitted
  - Low network load but high processing demands for the smartphone processor
- **Server-side**
  - the full image file is transmitted and processed in the server
  - High network load but low processing burden for the smartphone processor
- **Hybrid**
  - various combinations of the two strategies above.

Server side



Client side

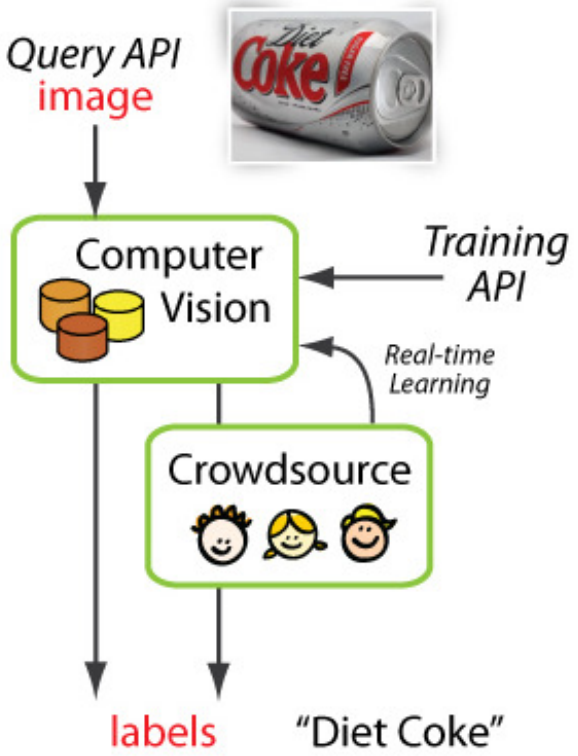


source: Liu et al., Mobile Visual Search, Linking Printed Documents to Digital Media, CVPR 2010, Demo

Technologies IV

Image recognition based on:

- **Nearest-neighbor**, where the query image is matched with one or more image(s) with similar content.
- **Object recognition**, where the query image is examined by all available models for the existence of certain object patterns.
- **Watermarking**, where encryption techniques are used to cast (and then detect) a content identifier inside the image pixels.
- **Human computation**, where human annotators are employed to recognize the image content when the automatic detection mechanism fails.



source:
http://www.iqengines.com/applications.php

Google Goggles

Use pictures to search the web.

Visual Image Search

OCR

Google

Logos, Bar codes

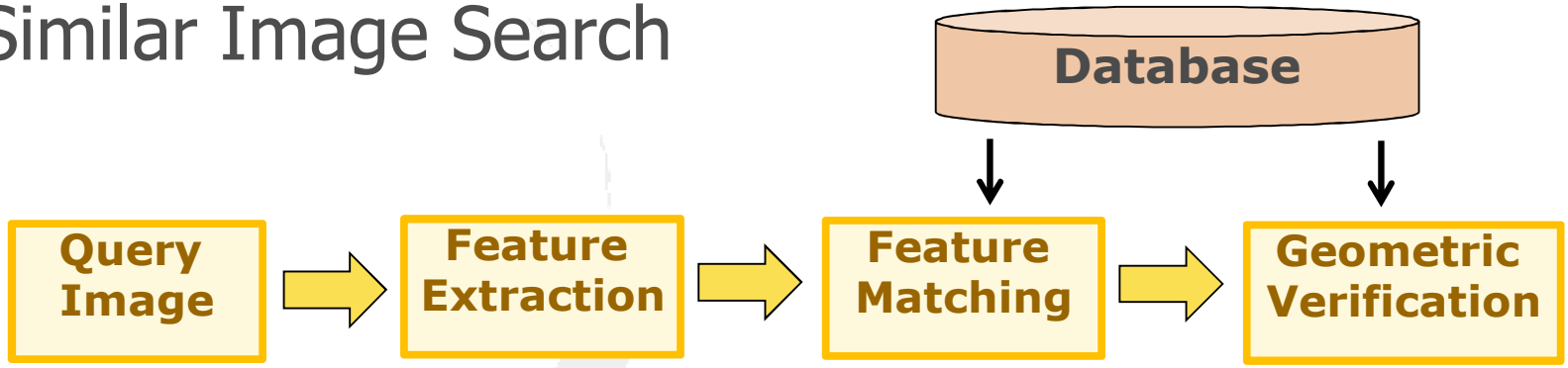
Wine

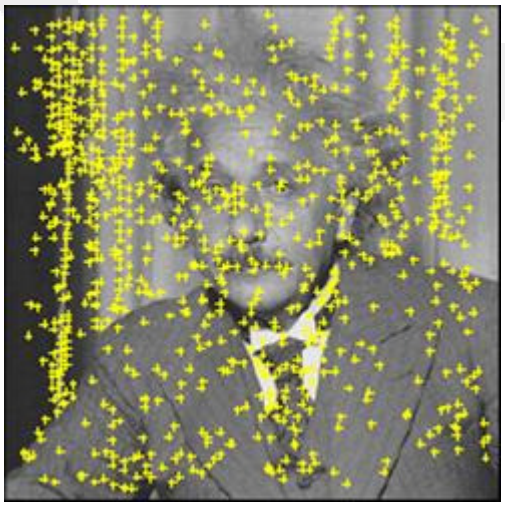
Custom search

Landmarks

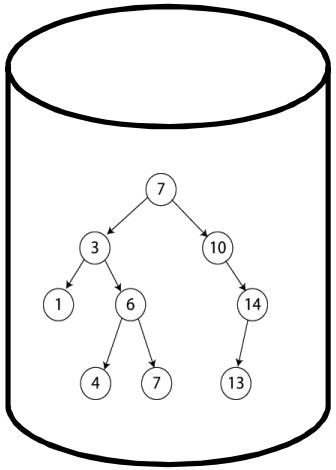
Generic search

Key-enabler technology: Similar Image Search

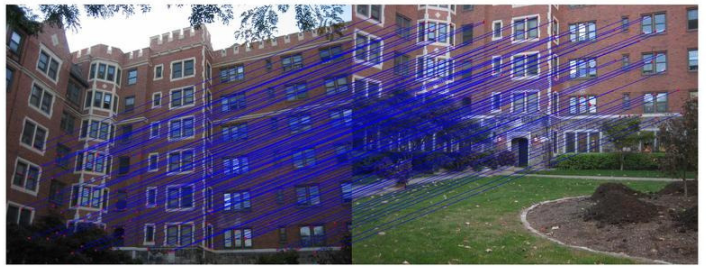




Features-> SIFT, SURF, Bag-of-visual-words

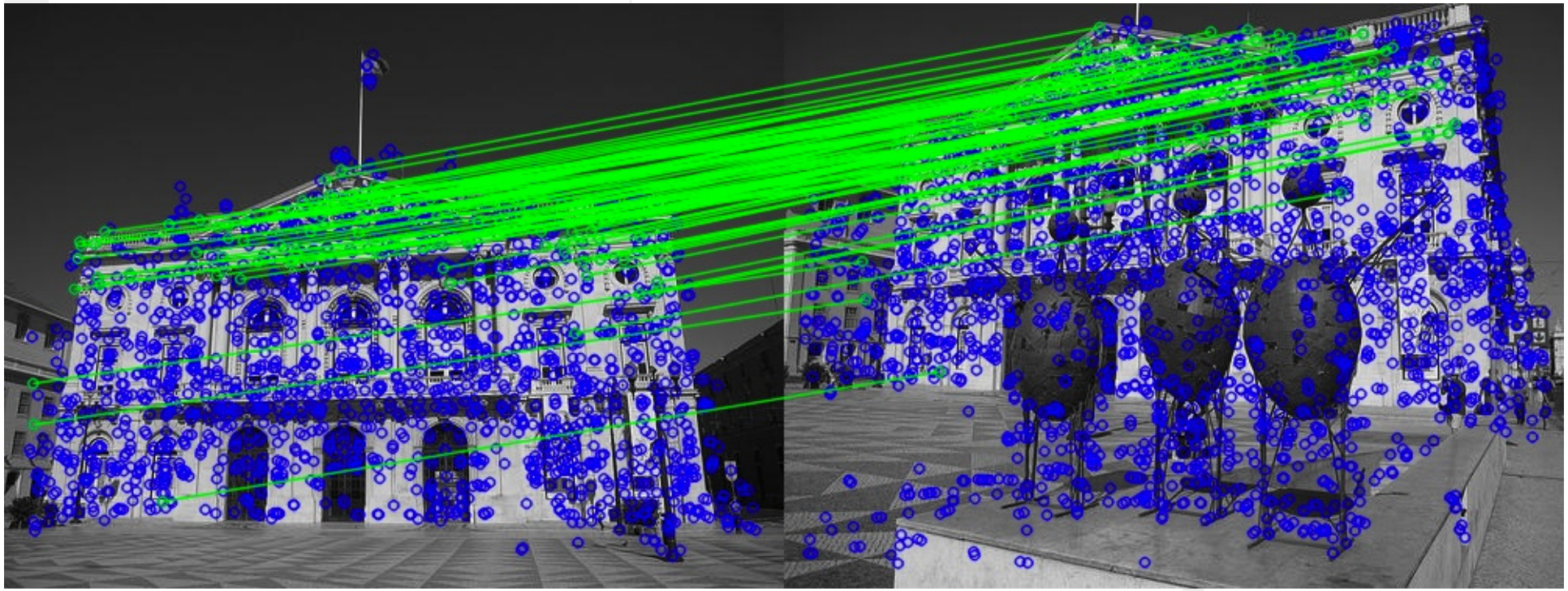


Indexing & Matching-> hierarchical clust., LSH, Tree structures



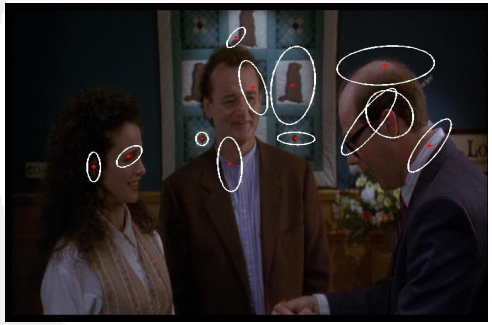
Geometric Verification-> RANSAC

Key-enabler technology: Similar Image Search

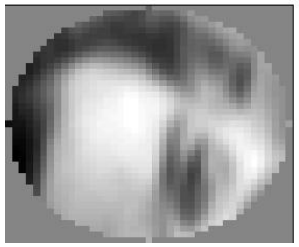


Key-enabler technology: Object recognition

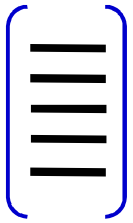
source: Rob Fergus, ICCV2009 presentation



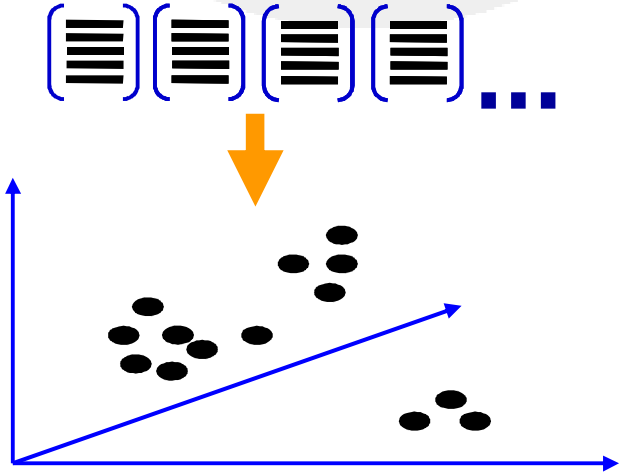
Detect patches



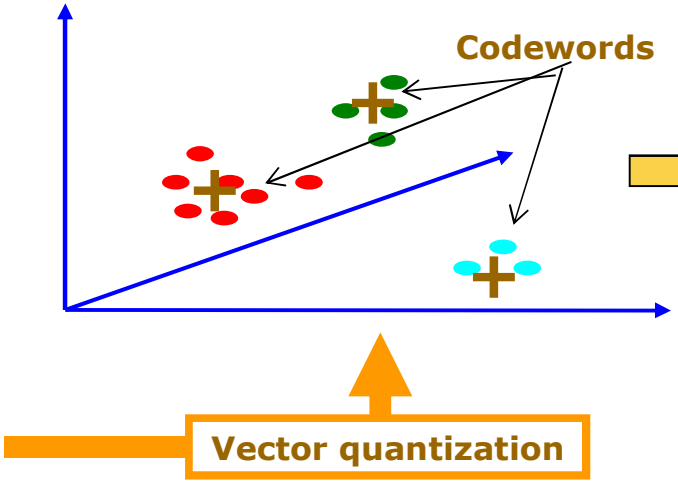
Extract patch



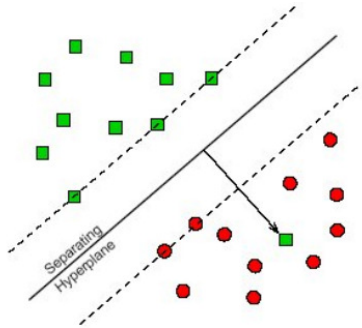
Compute descriptor, e.g. SIFT



128-D SIFT space

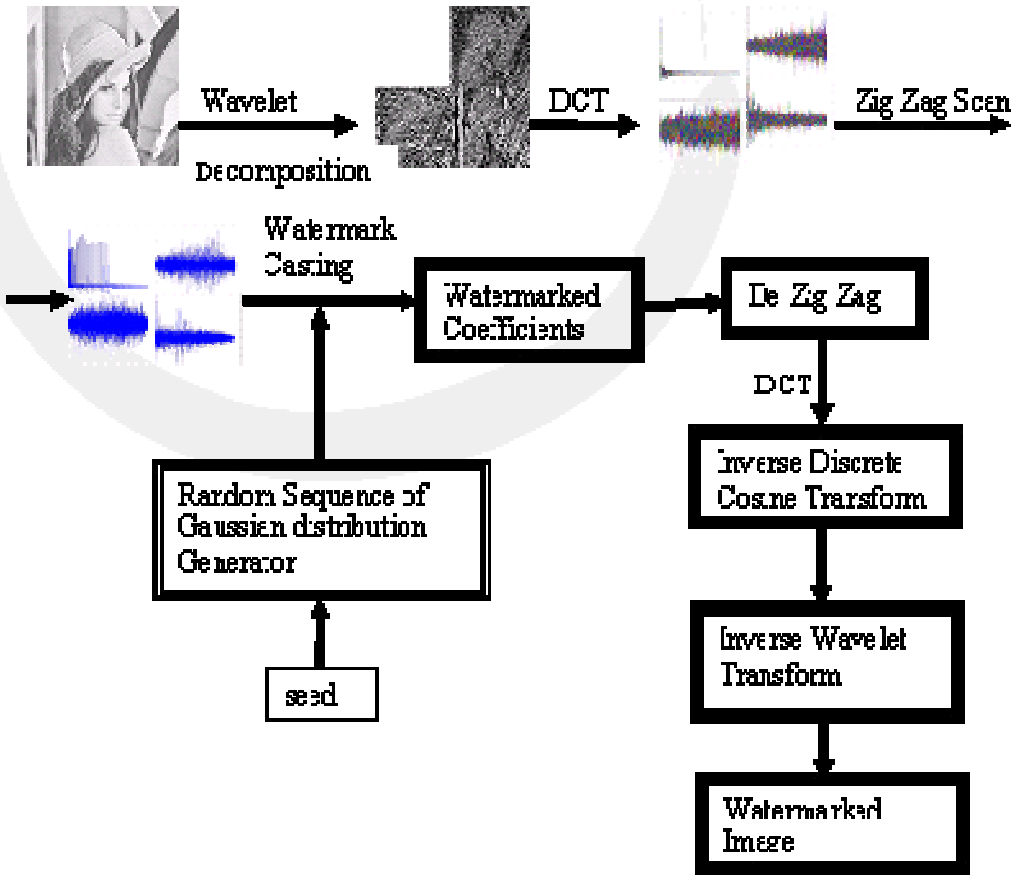


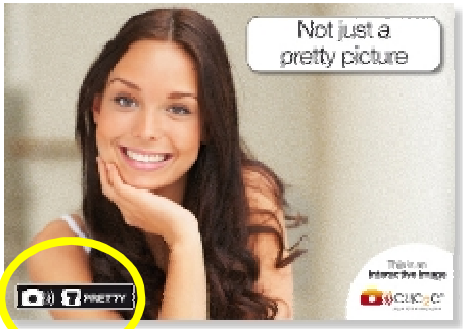
Vector quantization



SVM-based classification

Key-enabler technology: Watermarking



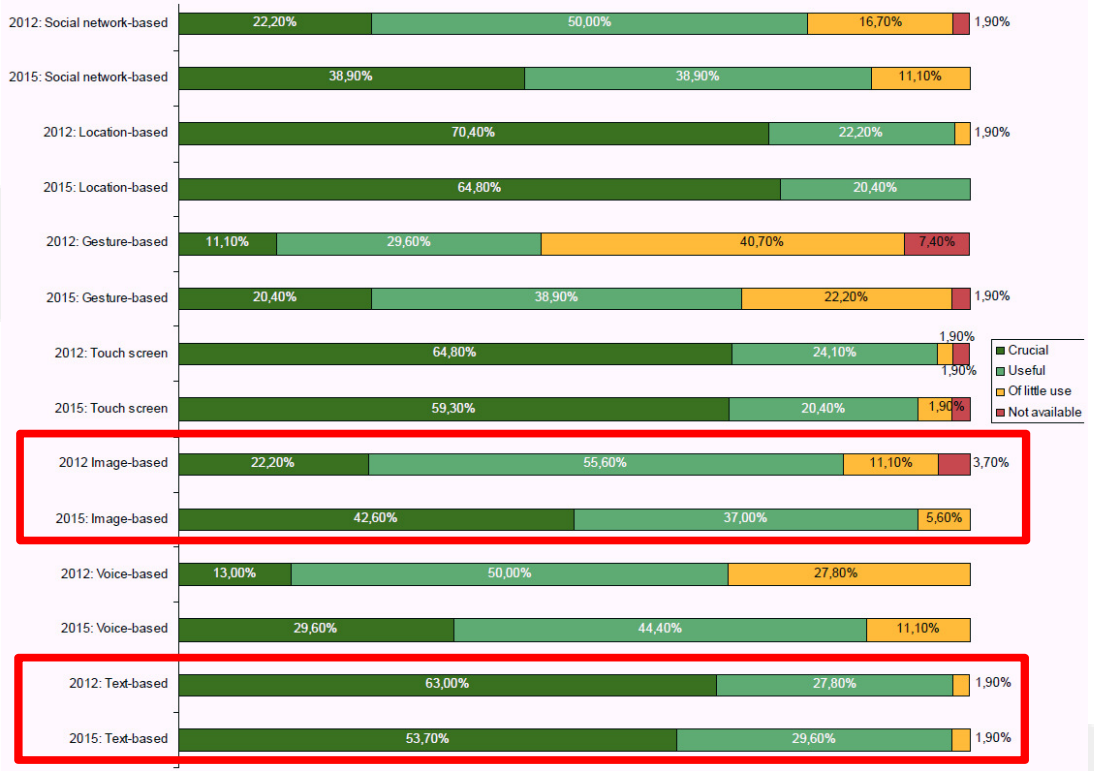


Subband-DCT watermarking

User needs and usage

Users use mobile search when:

- They are on the move
- The don't have access to a PC (e.g. waiting in the airport)
- User's location and surroundings is important for retrieving relevant information (e.g. find nearby gas stations)
- It is more convenient to use their phone (e.g. it takes longer to switch on their PC)



Expected interface usefulness for mob. image search in 2012/2015

- They need OCR capabilities for translating for instance, foreign billboards, traffic signs, etc.

Image-based interfaces are expected to acquire an equal important role with text-based interfaces

Business models

Intangible benefits model:

- Provide users with free services in exchange of their attention, loyalty and trust.
- At a later stage "monetize" this attention, loyalty and trust to gain profit.

Monetary benefits model:

- A transaction or a subscription process takes place and customers are required to pay in exchange of services or goods.

Three models are currently adopted by the majority of services:

- **Advertising-based model** , that uses visual search to find advertisements related to the content of the query image in a way similar to Google adds.
- **Software-as-a-Service (SaaS) model**, that relies on charging the customer for accessing or extending the image recognition-based, search engine.
- **Contract-based**, where an advertising company employs the core mobile image search functionality in the context of an advertising campaign.
- **Referral fees**, promoting products or services to new customers through referrals and collecting the respective fees.

Case study

Examined 12 services in terms of:

- The employed technology
- The target domain
- The adopted business model

Details about these services (applications) can be found in:
Chorus+, deliverable D4.1- Appendix "Study on Mobile Image Search", May 2011, http://www.ist-chorus.org/Mobile_Image_Search

Service comparison

SERVICE	TECHNOLOGY	TARGET USERS	BUSSINESS MODEL
Kooaba	Image recognition	Entertainment	SaaS-based
IQ Engines (oMoby)	Image recognition Crowdsourcing	Shopping	SaaS-based
Mobile Acuity	Image recognition	Consumers, Marketers	Contracting with marketers
LinkMe Mobile	Image and audio recognition	Consumers, Marketers	Contracting with marketers
Snaptell	Image recognition	Consumers, Marketers	Contracting with marketers
Point&Find (NOKIA)	Image recognition	Mobile users, Marketers	Engage more users in NOKIA experience
Gazopa	Image recognition	Shopping, Mobile users	Improving on-line shopping experience
Google Goggles	Image recognition	Mobile users	Advertising-based
Clic2c	Watermarking	Entertainment	Contracting with marketers
WeKnowIt IMG REC	Image recognition	Tourism	Touristic promotion actions
Wizup	Image and audio recognition	Consumers, Marketers	Contracting with marketers
TinEye Mobile (Snooth)	Image recognition	Wine industry	Advertising-based

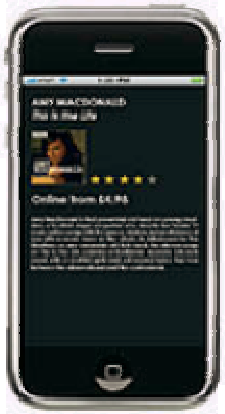
- **Image recognition** is the dominant technology employed by the majority of existing services.
- The most popular domain that is targeted by the existing services is **shopping**.
- Although contract-based seems to be dominant business model, we expect that sooner or later the **advertising-based** model will dominate this market in the same way with text-based search.

mobileacuity

Point
Point at a product, or an image of the product in an advertisement, using your mobile phone's camera

Decide
Receive product information including reviews, samples, online and local prices and availability

Buy
Make an immediate purchase on your mobile or buy later at home from the stored wish-list



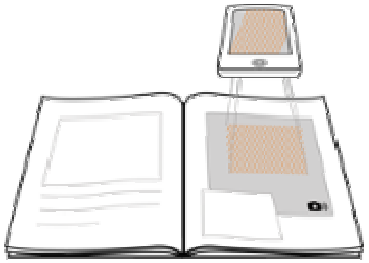
Business model:
Referral fees

Technology:
Image Recognition

Make your pictures interactive with MyCLIC2C®

Download the application

Discover images marked with this logo



Snap a photo of the image

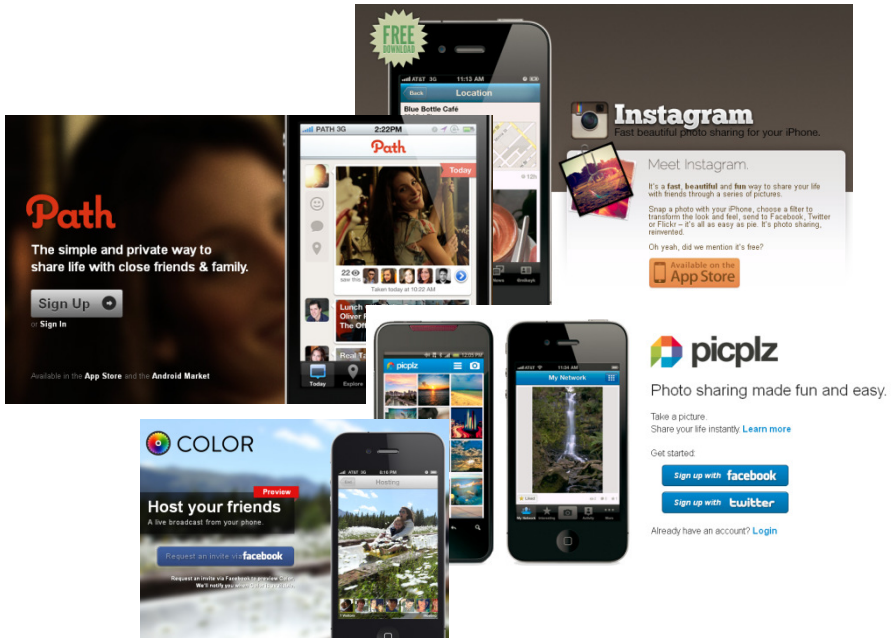
Enjoy the multimedia content

Business model:
Contract-based

Technology:
Watermarking

Future Trends

Photo sharing



Augmented Reality

Conclusions

- The robustness of the employed **image recognition technology** and the use of suitable **search interfaces** are perhaps the two key-elements for the success of mobile image search services.

- As the image recognition technology evolves, we expect that the majority of existing services will move away from the currently dominant, contract-based business model, and favor a **monetary benefits** approach primarily focused on **advertising**.

- The growing interest around **mobile photo sharing** and **AR applications**, will motivate the development of new innovative applications with mobile image search being the core functionality.

- Finally, the **rich context associated** with mobile image search queries, will definitely provoke the research community to exploit this context for **improving the performance** of current image recognition technologies.

Questions

?

Thank you!

http://mklab.iti.gr