Nevan Vision
Mobile Visual Search

Dr. Hartmut Neven, CTO
neven@nevenvision.com
Why Visual Search for Mobile Phones?

Camera Phones are an Enormous Market

- More cameras are bundled with mobile phones than sold standalone or bundled with any other device
- 365 million units shipped in 2005; 903 million units forecast to ship in 2010
- 87 percent of all handsets shipped in 2010 will be camera phones
- By 2010 camera phone users will have captured a total of 228 billion images
Neven Vision
Visual Search Market Activities

i-Scout™
From Mobile Marketing to Visual Search

Mobile Identifier™
In-Field Multi-Biometric Identification

Application Development
Photo Sorting, Visual Effects, etc.
i-Scout™
Visual Mobile Search

nevenvision
MACHINE VISION TECHNOLOGY
i-Scout
Visual Mobile Search

1. Snap a photo with a camera phone
2. Receive relevant content to your phone
i-Scout
Retrieving Mobile Content

1. Snap photo of a movie poster

2. Photo sent to Neven Vision servers

3. Neven Vision recognition engines identify poster image and find associated content

4. Content sent to phone (URL, text, sound, etc.)
Building a Visual Search System

User Requests

Data Processing/Recognition

Data Input & Ad Management Systems

Neven Vision Recognition & Data Servers

Campaigns | Content Searches | Shopping Tools | Blogs | Etc.

Users

Object | Face | Barcode | Recognition Engines | Character

Data Management/CRM

Advertiser Management

Marketers | Movie Partners | Music Partners | Shopping Engines | Public or Private Data

CRM Database | Billing

© 2005 Neven Vision • Company Confidential • Page 7
Brand and direct marketing with performance feedback
- Connect to customer at moment of interest and create lasting connection
- Rich, multi-channel, highly interactive, repeated/reinforcing impressions
- Customer action to initiate
- Get tracking data, marketing feedback

Fun, active, and highly convenient experience for consumers

Higher click rates than in comparable SMS campaigns
Image-Based Search: Mobile Advertising

Instantaneous and Interactive

1. Photo of object, advertisement, etc. is captured on camera phone.

2. Object is recognized with predetermined coordinates. Web interface provides information about the detected object.
Mobile Companion: Travel Guide

1. Picture of a place of interest
2. Read history of the site
   Find admission prices & tour times
   Schedule a tour
Image-Based Search: Product Inquiry
Catering to the Mobile Lifestyle

1. Picture of Car

2. Receive product information
   Customize your options
   Get a price quote
   Receive purchase rewards / feature upgrades
Mobile Companions:
Movies, Music and Games

1. Picture of movie poster
2. Read reviews
   - See showtimes and purchase tickets
   - Watch the trailer
   - Download ringtones
Mobile Identifier™
In-Field Multi-Biometric Identification
The Mobile Identifier platform is providing our officers with critical identity information at their fingertips... Every police officer needs to have one on his belt.

Commander Charlie Beck
Los Angeles Police Department

Camera-enabled Pocket-PC device captures an image of subject's face. Face is compared against faces in on-board, embedded database. Identity is verified.
Mobile Identifier™
In-Field Identification with Multi-Biometrics

- Mobile device for in-field identification using multi-biometrics
- Face and finger biometrics
- Version 2.0 integrates skin and iris recognition (SIMBA)
- Device holds up to 200,000 records locally
- Allows for wireless WAN connectivity
- Successfully field tested with numerous organizations
- High demand for Mobile Identifier in the law enforcement, military and corporate security markets.
- Acts as force multiplier
Secret Sauce: SIMBA™
Single-Image Multiple Biometric Analysis

- 2D high resolution images allow the seamless integration of **Facial Feature, Skin Texture** and **Iris** Analysis into a single recognition engine.

- Required resolution in pixels between the eyes:
  - Facial feature >25*
  - Skin texture >100*
  - Iris >600*  
  (6 Megapixel Camera)

- Neven Vision’s Facial Feature Tracking is the necessary base technology for SIMBA.

*Numbers assume a sharp image.
DoCoMo winter models 2005 are equipped with facial biometrics.
# Neven Vision Core Technology

## A Broad Set of Integrated Recognition Modules

<table>
<thead>
<tr>
<th>Module</th>
<th>Technology</th>
<th>Applications</th>
</tr>
</thead>
<tbody>
<tr>
<td>fD</td>
<td>Face Detection</td>
<td>photo sorting, automatic camera control: focus, zoom</td>
</tr>
<tr>
<td>ffD</td>
<td>Facial Feature Detection</td>
<td>facial gesture recognition, selection of optimal picture, redeye elimination,</td>
</tr>
<tr>
<td>ffT</td>
<td>Facial Feature Tracking</td>
<td>video special effects, character animation, video games, gaze tracking and drowsiness detection</td>
</tr>
<tr>
<td>fR</td>
<td>Face Recognition</td>
<td>physical and logical access control, surveillance systems, identity verification, photo sorting</td>
</tr>
<tr>
<td>oR</td>
<td>Object recognition</td>
<td>visual mobile search, security applications include tattoo, stolen artwork and driver license recognition</td>
</tr>
<tr>
<td>BCR</td>
<td>Bar Code Recognition</td>
<td>visual mobile search, contains licensed technology</td>
</tr>
<tr>
<td>OCR</td>
<td>Character Recognition</td>
<td>visual mobile search, contains licensed technology</td>
</tr>
</tbody>
</table>
Overview of Recognition Engine

**Image**
- Points of Interest Detector (optional)
- Extraction of Feature Sets

**Database of Learned Objects**
- Feature Sets Representing Stored Objects (ideally parameterized)

**Matching**
- Select an initial set of elastic graph configurations and parameter settings that transform feature sets
  - Similarity Engine
  - Translation Engine
  - Parameter Engines
  - Controller for elastic graph moves and parameter settings that transform features

**Compare maximal similarity score to a predefined threshold to decide whether an object is visible in an image.**
Interest Point Detection
Choice of Features

- Multiple feature types are employed concurrently
- A corner stone feature are Gabor Wavelets
- Motivation for the use of Gabor Wavelets
  - Good experimental results
  - Favorable signal theoretic properties
    - Optimal localization in space and frequency domain
    - Amplitude is invariant under small translations
    - Phase information for precise localization
  - Biologically plausible
Finding of Corresponding Points

1. How similar is the feature $\tilde{f}(\bar{x}^c)$ extracted at $\bar{x}^c$ to a set of sample features $\{\tilde{f}^G(p_i)\}$?
2. How different is $\vec{x}^G$ and $\bar{x}^c$?
3. Which parameters $\vec{p}$ characterize $\tilde{f}(\bar{x}^c) = \tilde{f}(\bar{x}^c, \vec{p})$?

Only Neven Vision employs dedicated engines to address questions 2) and 3).
A neural network based translation engine proved superior over phase based translation engine but is less general.
Experimental Setup to Study Feature Manifolds

Light Stage 2.0 by Paul Debevec

3 Dof Robotic Arm by Benahm Salemi
Parameter Engine

Using parameterized feature sets is key to achieving illumination and pose invariance.
Camera phones constitute a large and fast growing market.

Image recognition forms the basis of new services that will become pervasive.

Machine vision enables broad range of services: from identity management via visual search to entertainment applications.