Computer Vision Now and Then: A Personal Journey

Gérard Medioni

Overview

- Bio sketch
- Research outside academia
- ***Childhood*** (the 90s): OptiCopy, EPSIS, lessons
- ***Adolescence***: Geometrix, Bigstage, lessons
- ***Adulthood***: PrimeSense, lessons
- Today's landscape

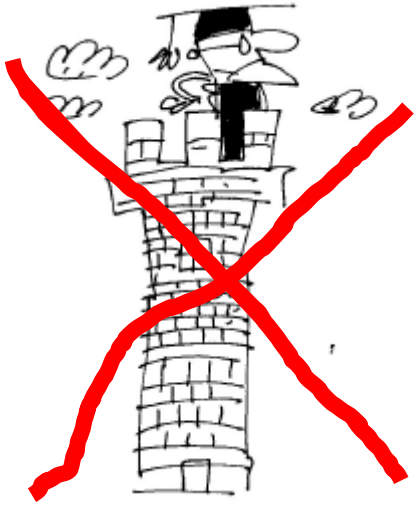
Bio Sketch

- Engineering 1st degree: Telecom Paris ‘77
- USC “Inbred”: PhD 83
- USC CS Department Chairman: 01-07
- Broad range of research interests
- Conference organizer: CVPR, ICCV, WACV, ICPR, IWCV
- Strong interaction with industry

Research outside academia

- Not Ivory tower research:

  Invent a problem, solve it, ask if anyone cares…



- Valuable interaction
  - Good source of problems
  - Good resource for **academic** publications:
    - Identify core, not engineering, issues
    - Produces results outside original framework

Disruptive technology

- The good news is that it is disruptive, the bad news is that … it is disruptive

- It opens road to new opportunities, but
- Does not fit into existing models, =>hurdles to adoption

Childhood (the 90s)

"Using Computer Vision in Real Applications: Two Success Stories"

at the MVA conference in 1996 in Tokyo.

- Registration of half-tone separations (Opti-Copy)
- Real-Time insertion of content into a video (EPSIS)

Definition of “success”

- Build a gizmo that does what is was designed to do…
- Technology success
- Business side coming up…

My comments in ‘96

- *The field of computer vision has matured, and many techniques are producing consistent results*
- *While the techniques are understood, we still do not have a bag of tools which we could use as plug-in components of a solution*
- *Instead, each application requires a large amount of effort to customize algorithms*

Opti-Copy

Automatic Registration of Halftone separations

Core: Efficient edge detection, robust correspondences, high precision matching

Tech achievement

- 10 sets/hr vs. 4 for humans
- By the time was ready to be adopted,

the industry switched to digital, making the technology obsolete

Lesson 1

Keep aware of the global context and anticipate disruptive changes which may disrupt your disruptive technology.

Knowing your current competitive landscape is not enough.

EPSIS

- Real-Time technology for seamless insertion of computer-actuated content into a video
- Virtual Advertising
- Sports Enhancements
- Live or post-production
- Example

Core: Motion stabilization, panorama, feature tracking, object detection, image matting, lighting compensation, AR

Achievements

- Technology worked well
- Deployed for several high profile sports events

- Potential for huge business opportunity
- $$ from advertisers

- But 3 companies failed in this endeavor ?!

Roots of this business failure

- Disruptive
  - How to price?
  - How to insert in the chain?
- Too many parties to the deal
- Today:
  - Post-processing mostly
  - Low cost service similar to “slow-motion”

Lesson 2

Technology without a business model is doomed from the start

Technology and business model need to proceed in parallel

Adolescence(- 2008)

- Geometrix
  - 3D face modeling and recognition
  - 3D virtual try-on of glasses
- Bigstage
  - Insertion and animation of 3D face in video

My comments (2007)

My 1996 paper receives in 2007
"most influential paper of the decade award":

- *Computer Vision is now a mature field*
  - *Set of techniques that work*
  - *Implementations available (OpenCV, SourceForge, Matlab, …)*
  - *Datasets for validation/comparison*
- *Cameras are cheap*
- *Interfaces are standard (USB, Firewire,…)*
- *Both storage and computing continue to follow Moore's law*

Geometrix

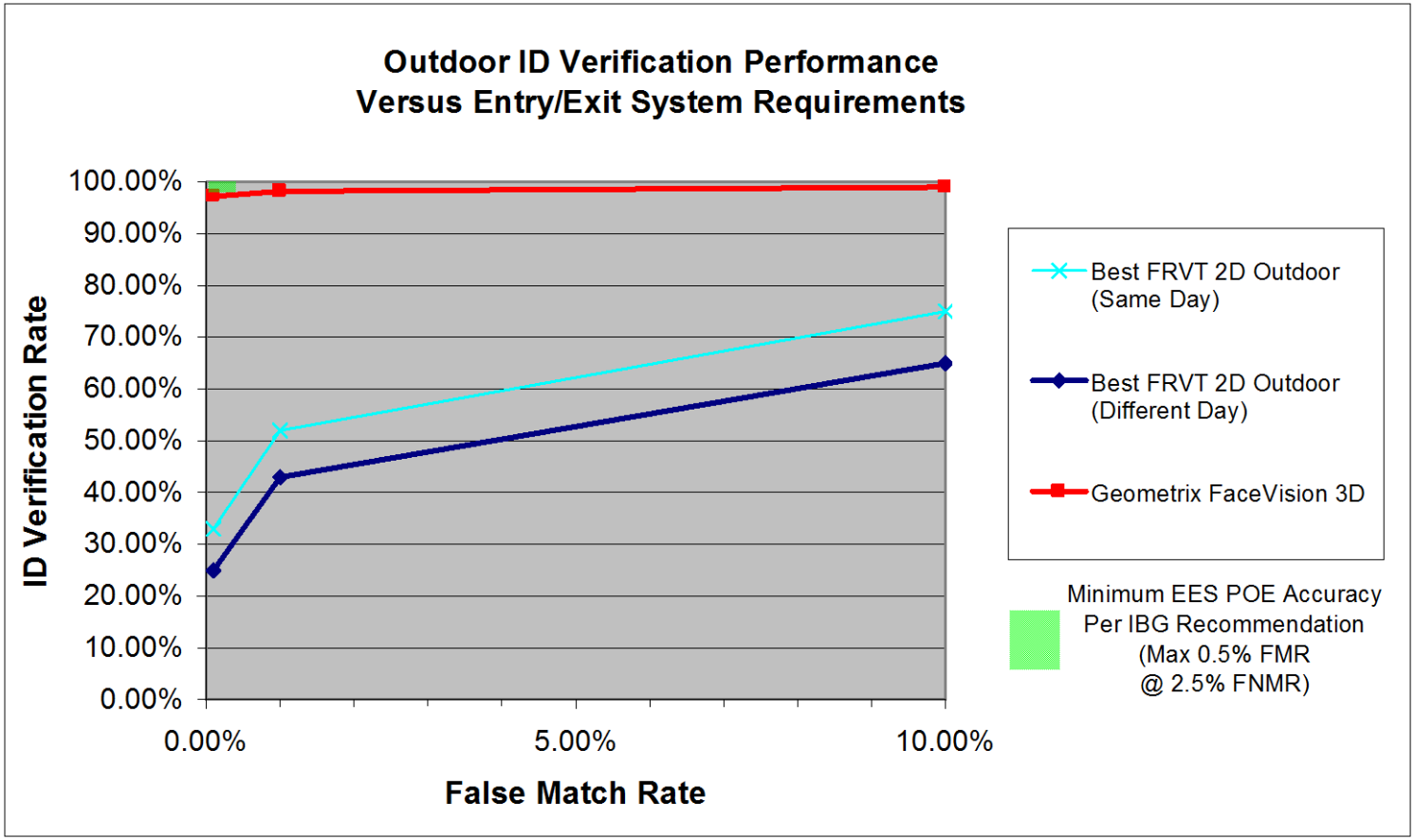
- CTO
- Robust Automated Face Modeling and Recognition Based on 3D Shape
- Deployed in Cobb County Adult Detention Center for booking/releasing inmates

Core: Stereo matching, 3-D modeling, 3-D robust alignment, 3-D face recognition, virtual try-on

Approach

- 3D Sensing
  - Passive sensing
  - 3-D inference from images
  - Low-cost COTS components

Authentication Performance



Business

- Acknowledged superior performance for 3D face recognition
- Company survived a few years on government research contracts, then folded
- Tried different markets (jails, optical stores)
- Required new infrastructure for deployment, but … system was **prototype**

Lessons 3 and 4

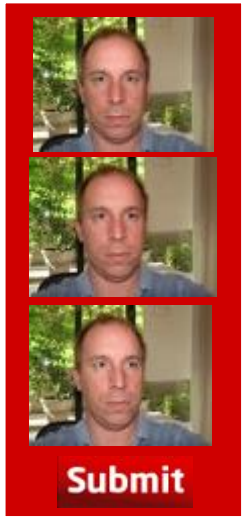
Focus is essential for a small company

A product is miles away from a prototype

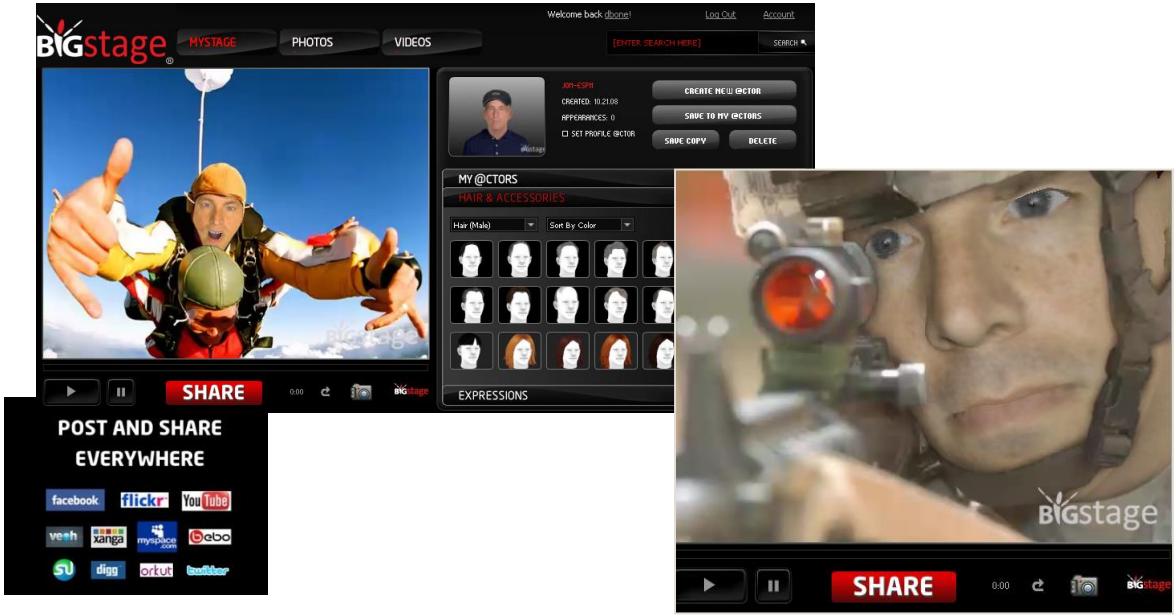
Bigstage

- Create and host a high fidelity 3D version of you, that can be personalized and inserted into online media and activities – instantly!

Take 1 to 3 pictures



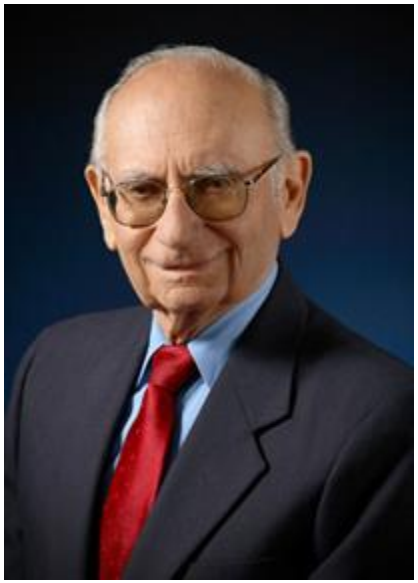
Personalize, play, comment, and share with friends



Core: 3-D face and head modeling, expression modeling, lighting correction, face animation

Example

- Prof George Bekey's 75th birthday and roast



Business

- Company was raising B round in 2008..
- Ran out of money
- Never had a business plan articulated
- Tried to monetize before finishing product

Lesson 5

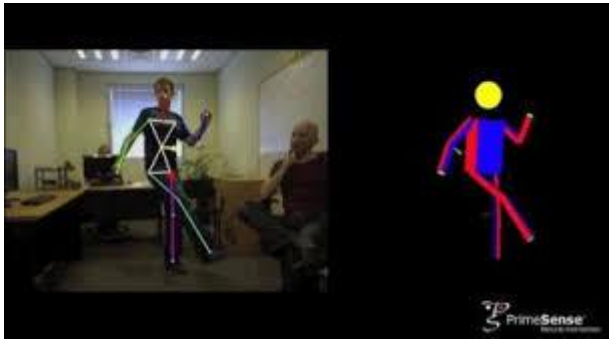
- Previous lessons apply, and

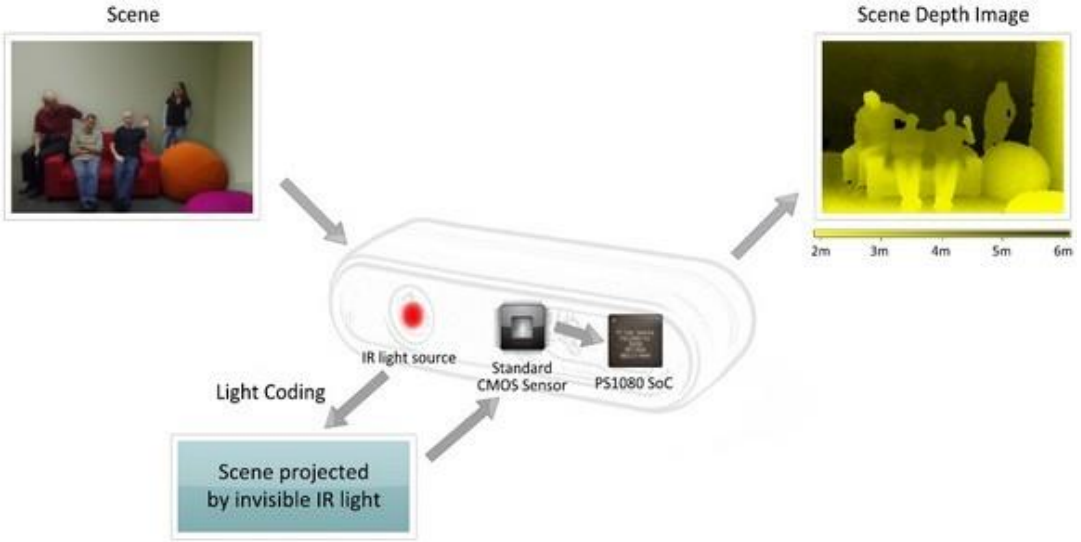
Timing is a predominant factor for success

Adulthood (- today)

- Primesense
  - Consumer grade 3D image capture for ~$100
  - Body as a User Interface
  - Gaming applications
  - Disruptive success
    - Revolutionized several fields (robotics, …)

Primesense





Core: Stereo matching, active ranging, 3-D body modeling, motion capture, gesture recognition, …

Business

- Company grew to >200 people
- Microsoft major customer (Kinect-1)
- Sold to Apple in Dec 2013

Lessons

- **Bold vision**
- **High risk proposition**
- **Strong leadership both tech & business**
- **Never deviated from business model**
- **Ability to adapt to ambiguity and change**

Landscape today

- Open Source tools (Dlib, OpenCV, …)
- Annotated datasets (ImageNet, …)
- Emergence of Deep Networks
  - Environments (Caffe, Torch, TensorFlow, …)
  - Pre-trained networks
  - Data vs algorithms

Landscape (cont.)

- Open access to code
- Scientific publishing revolution
  (ArXiV, Open access, online publishing)
- Frenetic pace of progress update

Are we done?

- No, far from it!
- Scientific challenges still remain
  - Shallow understanding of deep networks
  - True semantic interpretation still open
  - Video Analysis challenging

Nevertheless

- Computer Vision that works to solve real problems in commercial applications
- Sandbox for Science+Engineering
- Many opportunities for
  - Startups
  - Large companies make significant investment in CV (Amazon, Google, Apple, Microsoft, Facebook, …)

We live the Golden Age of Computer Vision