IC Technology for Cultural Heritage

Alberto del Bimbo
Professor
Università di Firenze
Museums, Cultural Heritage sites exist to expand our knowledge, stimulate our senses, expose us to new experiences and engage our participation, as we explore our shared history and the natural world.

IC technology has expanded the dimensions of Museums and Cultural Heritage sites. Turning visitors from passive observers to actively engaged participants. Inspiring stakeholders to supplement static display techniques with dynamic, responsive, participatory environments.

From object-centric to human-centric IC technology. Delivering a user-centered personalized dialog between the site and its visitors.

Cultural heritage sites as intelligent information spaces that engage visitors while being aware of their age, genre, behavior and personal interests…
Consider technological trends
Mobile technology

Going mobile
Mobile is the principal media of access to the internet

Source: Social Stand Media (2014)
Consider technological trends
Social data analysis

Cloud Computing and Big Data Analytics

Increasing convergence of innovative technologies Mobile, Social, Cloud, Big Data, Internet of Things..... into applications and services centered on the user experience

The analysis from social data typically discovers macro-interests that represent habits of people
Continuous and Ambient User Experience

Mobile apps will remain important but there will be a shift on providing an experience that exploits different devices, including IoT sensors, common objects...

The experience will seamlessly flow across a shifting set of devices and interaction channels blending physical and virtual environments.

Look for augmented reality to thrive in unexpected places....
Consider technological trends
“Intelligence” in machines

**Machine Learning, Deep Neural Networks**

Move beyond classic computing to create biologically inspired systems that can autonomously learn to perceive the world on their own.

They ground on the availability of large masses of data to address key challenges in **automatic visual understanding**, **automatic speech and audio recognition**, **natural language processing**,.....
Consider technological trends
Computer vision enabling

**Computer Vision**

A powerful artificial sense to extract information from images: about places, objects, people....

Capable of:
- performing recognition of objects, people irrespective of difference in appearance, contexts....
- answering visual questions...
- understanding behaviors and situational conditions of people....

Either from fixed or mobile cameras
NEMECH New Media for Cultural Heritage
Competence Center Univ Firenze - Tuscany Regional Gov

Established on 30/06/2011
Approval 2779 Tuscany Regional Government
Educational activity
Research and research transfer activity

**The Hero project**  
Academia Museum, Florence, 2015
The Bahnlandschaften project
Forte di Fortezza, Bolzano, 2015-2016

MSKinect depth sensor
NEMECH New Media for Cultural Heritage
Competence Center Univ Firenze - Tuscany Regional Gov

The Galleria Vittorio Emanuele project
Military monument Cima del Grappa, Udine, 2016
(not realized)

Sensorized ceramic floor
4.5 mm thickness
1 sensor every 12 cm²
Italian Architecture Society meeting
Leopolda Station, Florence, 2016

Expanded reality

Heliodisplay holographic technology
The MNEMOSYNE project
Museum, Florence, 2013 - 2016

Bargello

1 interactive table
4 cameras
5 computers: 1 MASTER, 4 Computer Vision SLAVES
The “intelligent audio guide” project
Wearable Computer Vision technology, 2016-2017

Deep Network-based computer vision
on NVIDIA Jetson Tegra K1
system-on-chip wearable processing board

Fig. 2. Feature extraction procedure for an artwork detection on a single convolutional feature map.

Fig. 3. Our network architecture, with tensor size and layer numbering.

the discriminative power, we concatenate such descriptor computed on two different feature maps. The region is remapped from the frame to the convolutional activation map with a simple similarity transformation.

Considering the activation map of the $n$th convolutional layer, we have a tensor dimension of $W_n \times H_n \times C_n$. After the reprojection of the bounding box onto the feature map we end up with a smaller tensor with a size of $W_{n,bb} \times H_{n,bb} \times C_n$; $W_{n,bb}$ and $H_{n,bb}$ depend both on the network layer and the bounding box geometry, while $C_n$ depends solely on the network layer and represents its number of channels. The max-pooling operation of the $C_n$ channels over the $W_{n,bb} \times H_{n,bb}$ values generate a feature vector that is independent from the dimension of the bounding box.

Considering the architecture in Fig. 3 one could wonder which features are best to recognize the specific framed artwork, since leftmost layers have higher resolution and mostly represent the low-level structure of the image, while rightmost ones, are low resolution but encode higher level information, closer to the image semantics.

After an experimental evaluation, which is detailed in Sect. 7, we selected, as combination, the features from layers 3 and 4, yielding a feature of size 768. The final bounding box descriptor is obtained by concatenation of the two max-pooled regions values and is $L_2$-normalized.

Considering a pre-acquired dataset of artwork patches $p_i \in D$ and their artwork labels $y_i$, for each detected artwork $d$ we predict a specific artwork label $\hat{y}_p$ finding the $L_2$-normalized.
The “intelligent audio guide” project
Wearable Computer Vision technology, 2016-2017

SIFT-based computer vision
on NVIDIA Jetson Tegra K1
system-on-chip wearable processing board

L. Seidenari, C. Baecchi, T. Uricchio, A. Ferracani, M. Bertini, A. Del Bimbo
EVA International Conference, 2017
The SOCIAL MUSEUM AND SMART TOURISM project
MIUR National Cluster project 2016-2018

Project partners
UNIVERSITA' di FIRENZE,
UNIVERSITA' di ROMA LA SAPIENZA,
IUAV
ENGINEERING SpA,
VITROCISET SpA,
RAI SpA,
SMI CONSORTIUM
People at MICC UNIFI