



VISITO Tuscany & Visual Recognition



Stefano La Terra Inghilterra

VISITO - overview

VISITO – acronimo di **VI**sual **S**upport to **I**nteractive **TO**urism – è un'applicazione per smartphone sviluppata nel 2009 da un gruppo eterogeneo di istituti e imprese private¹, finanziato dalla Regione Toscana. Il suo scopo è migliorare l'esperienza culturale dei turisti che visitano le città d'arte toscane, grazie alla capacità di riconoscere luoghi e opere d'arte.

VISITO svolge questa funzione attraverso l'uso della fotocamera presente sui moderni smartphone: il turista non deve far altro che scattare una foto a ciò che lo interessa – sia esso una statua, un palazzo, un elemento architettonico – per ricevere sul proprio schermo tutte le informazioni in proposito. Potrà così apprendere di più su ciò che ha di fronte, nonché accedere ad approfondimenti sull'autore dell'opera ed altri argomenti correlati.

Durante la fruizione dell'app mobile, VISITO rimane connesso ai propri server per consentire l'identificazione delle foto scattate e ricevere i dati relativi, in forma di ulteriori foto, dati storici, e informazioni contestuali. Offre anche una Google Map con visualizzati i POI (Points of Interest) nelle vicinanze, sfruttando il posizionamento satellitare per un'esperienza culturale di massimo livello.

VISITO inoltre - sulla propria interfaccia desktop - consente all'utente di creare itinerari utilizzando il proprio database di luoghi d'interesse, opportunamente visualizzati su una mappa navigabile di Google Maps; non solo, in base alle passate esperienze dell'utente VISITO è in grado di suggerire ulteriori mete che potrebbero interessarlo. Non mancano i cosiddetti "itinerari istituzionali", che raccolgono le mete d'interesse principali per ogni città visitabile. VISITO fornisce infine la possibilità di effettuare la ricerca libera per parole chiave (es "chiesa", "battistero", ecc.).

Dopo la visita, infine, l'utente potrà rivedere le foto scattate durante il viaggio, in qualche caso anche sovrapposte a modelli 3D di piazze, e condividere con altre persone l'esperienza vissuta.

La proposta di VISITO è rivolta a diverse fasce di utenza – oltre a quella del soggetto turista - ciascuna delle quali ne ricava vantaggi non indifferenti:

"[...]

- Pubblica Amministrazione: VISITO-Tuscany costituisce uno strumento per la promozione e il
 marketing territoriale. Un Comune può ricorrere a VISITO-Tuscany per incentivare i turisti a
 visitare la città: VISITO può raccontare itinerari, invitare alla visita di mostre ed eventi, a
 valorizzare le arti ed i mestieri locali. Inoltre, VISITO è uno strumento molto efficace di
 comunicazione locale (cittadini e abitanti dei comuni limitrofi) per pubblicare manifestazioni ed
 eventi.
- Operatori territoriali (APT, infopoint): VISITO può supportare gli operatori nel servizio di assistenza informativa ai turisti, comunicare meglio, comunicare di più, ridurre le code e aumentare la soddisfazione dei visitatori.
- Imprese connesse al turismo (tour operators, compagnie aeree, associazioni culturali locali, operatori alberghieri, attività commerciali): VISITO consentirà di valorizzazione le destinazioni dei turisti tramite un'informazione aggiornata, personalizzata e collegabile agli esercizi commerciali.

[...]" (Sam Habibi Minelli, 2011)

¹ CNR di Pisa, ISTI, Alinari 24ORE, 3logic, Iperborea

VISITO – interaction details

Mobile interface

L'interfaccia mobile è suddivisa in quattro sezioni principali: visite, caccia al tesoro, camera e mappa, accessibili tramite un menù a tab orizzontale.

Visite

La sezione visite si articola in visite suggerite e create dall'utente, visualizzabili alternativamente con uno swipe a destra e sinistra o mediante pulsanti dedicati. Entrambe le categorie vengono caricate dinamicamente tramite connessione dati, in tal senso VISITO non può operare offline.

La categoria creata dall'utente, in particolare, prevede un accesso remoto con account da crearsi sul sito web di VISITO.

Ciascun percorso di visita è visualizzato come un insieme di luoghi evidenziati (Points of Interest) su una mappa di Google; ciascun POI è interattivo e se cliccato fa accedere ad una descrizione corredata da galleria di immagini, e in alcuni casi da approfondimenti accessibili tramite il tasto menù dello smartphone che stiamo utilizzando.

- Caccia al Tesoro

La sezione Caccia al Tesoro è molto simile a quella visite, ma prevede un percorso a tappe dove la meta successiva viene suggerita con un indovinello quando si fotografa correttamente quella attuale. Ogni tappa è corredata da una scheda descrittiva analoga a quella delle visite tradizionali, ma aggiunge un pulsante "mappa" per localizzarne la posizione rispetto a dove ci troviamo e due pulsanti di navigazione per muoversi tra le tappe.

- Camera

Accedendo alla fotocamera dal tab dedicato possiamo fotografare un edificio, una statua o un qualsiasi soggetto d'interesse. I server di VISITO riceveranno la nostra immagine e, se presente sul loro database, sul nostro cellulare apparirà la scheda descrittiva con foto che abbiamo avuto modo di usare nelle sezioni precedenti. Nel caso in cui un riconoscimento certo al 100% non sia possibile, ci verrà comunque suggerito un elenco di luoghi d'interesse vicini a noi (via GPS).

- Mappa

Una mappa di Google con funzione GPS che ci localizza e mostra le mete d'interesse attorno a noi. La mappa è navigabile, pertanto possiamo cercare mete da raggiungere esplorandola e spostandoci virtualmente sul territorio italiano. Tutti i POI del database di VISITO vengono comunque caricati all'accesso, quindi possiamo in ogni momento cliccare su uno di essi per accedere alla scheda descrittiva.

Desktop interface

Il sito web dedicato offre all'utente una serie di funzionalità interessanti: pianificazione e rivisitazione a posteriori di itinerari, condivisione delle proprie foto, esplorazione 3D di alcune località come Piazza dei Cavalieri a Pisa e il centro storico di San Gimignano. In particolare quest'ultima esperienza sfrutta la tecnologia Photocloud sviluppata dall'ISTI, in grado di sovrapporre alla ricostruzione tridimensionale dei luoghi foto prospetticamente corrette a seconda del punto d'osservazione.



Figura 1: Photocloud visualizza Piazza dei Cavalieri.

Le foto scattate si possono organizzare comodamente per album, associare con descrizioni e, per chi ha l'animo più social, condividere su Facebook, Twitter e Del.icio.us.

Gli itinerari programmati da desktop sono automaticamente caricati sullo smartphone quando accediamo all'app mobile di VISITO, e sempre su desktop è possibile visualizzare i Point Of Interest in diversi modi:

- Evidenziati visualmente sulla Google Map
- Raggruppati per categoria o popolarità
- Ordinati secondo l'appartenenza a itinerari organizzati da esperti



Figura 2: VISITO Desktop: visualizzazione dei POI

What drives the functionality

La tecnologia che consente di individuare il soggetto di una foto basandosi sulle sue caratteristiche – dette "features" – si chiama visual recognition e può essere applicata nei campi più svariati come la traduzione del testo, la navigazione satellitare, la realtà aumentata etc.

È grazie ad essa che l'User Experience raggiunge un nuovo livello, e un'ottima spiegazione del perché è data dal seguente estratto della conferenza dell'operatore Telefónica al Mobile World Congress del 2010:

"GPS antennas have an error of several meters (depending on many factors, such as weather or sky visibility). Such error introduced into the point of view of the user is non-relevant for distant points of interest (POI). What happens with closely located places is another story. An error of 2 or 3 meters is enough for the display to show information about a POI behind us, just in the middle of the screen (i.e. as if it was in front of us). Imagine we are looking for a review of a restaurant that is in front of us, the question is: should we be seeing the review of another restaurant one block away? Here is where the user experience breaks." – (David Marimon, 2010)

In pratica, per gli scopi della realtà aumentata la precisione offerta dal GPS non è sufficiente – l'errore medio di posizionamento è nell'ordine di qualche metro. Combinata con il riconoscimento visuale è tuttavia utile restringere l'ambito della ricerca, e mostrare le informazioni rilevanti e in breve tempo; la localizzazione geografica infatti consente di restringere l'ambito della ricerca dei possibili risultati candidati a quelli vicini all'utente.

Va inoltre sottolineata l'interazione naturale con l'ambiente offerta dalla VR:

"[...]firstly, because the information or services provided are specific to THAT object that we are interested in; secondly, because it skips the sometimes cumbersome step of typing text on a mobile device." – (David Marimon, 2010)

Otteniamo nformazioni specifiche, praticamente a portata dell'occhio della fotocamera, che diventa il nostro occhio potenziato. Inoltre siamo in grado di effettuare la nostra ricerca in situazioni altrimenti difficoltose, dove una ricerca testuale sarebbe poco pratica: all'aperto, dove la visibilità dei display è ridotta; in luoghi affollati, dove scrivere è difficile; in movimento, dove non ne avremmo il tempo, ecc. La tecnologia VR non è nata offrendo la performance attuale (vedremo più avanti un accenno ad uno dei primi esperimenti), ma c'è stato un progresso tecnologico impressionante:

"I quickly discovered that this survey would be extremely difficult and virtually useless, as not a single person I initially attempted to survey had heard of either software, including those with phones capable of the Google app" - (Colleen, 2010)

In questo frammento di un articolo del 2010, l'autrice descrive la sua esperienza nel tentativo di raccogliere opinioni sull'app per smartphone Google Goggles, che consente il riconoscimento di oggetti (modelli di cellulare, copertine di libri, testo, etc.) utilizzando la fotocamera degli smartphone: concludere il sondaggio si è rivelato impossibile, poiché la maggior parte degli intervistati non era neppure a conoscenza di Goggles; non sapevano cosa fosse. Solo tre anni dopo, ecco l'opinione di Paul Dunay, esperto di marketing presso l'azienda specializzata Maxymiser:

"What I'm talking about is giving consumers control, a tool that allows them to quickly and easily learn something about a specific product or service. I'm talking using our smartphones to create intuitive gateways that bridge the real world with related digital experiences. Whether you call it Web 3.0 or the "internet of things," this is what mobile is all about and I believe visual recognition technology will play a key role in making it happen." - (Dunay, 2013)

L'internet delle cose, con l'utente interattivo al centro di esso, è diventato il fattore chiave del mobile, e un ruolo fondamentale verrà svolto dalle tecnologie di riconoscimento visivo. Dunay evidenzia come tali tecnologie siano in grado di fare da ponte tra il mondo reale e l'esperienza virtuale, arricchendo il primo attraverso la seconda in modo estremamente intuitivo.

È giusto ricordare che esistono approcci alternativi per il riconoscimento di oggetti nel reale, per citarne alcuni: *QR Code, Barcode, tag NFC* e *RFID*.

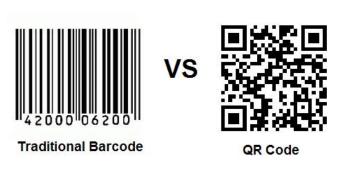


Figura 3: Barcode e QR Code

Si tratta di marker da apporre sull'oggetto d'interesse, in particolare i primi due necessitano di contatto visivo con la fotocamera, mentre *NFC* e *RFID* fanno uso di appositi sensori per la lettura. La necessità di usare dei marker presenta problemi in diverse situazioni, ad esempio su superfici irregolari che ne alterano il disegno o quando si va a coprire parzialmente delle immagini.



Figura 4: Da sinistra, un tag RFID e uno NFC

Inoltre questi marker funzionano bene sulle corte e cortissime distanze, ma sono poco adatti nel caso in cui non ci si possa avvicinare molto al soggetto, o si voglia identificare qualcosa da distante: basti pensare ad una stanza di museo molto affollata per il primo caso, o ad un edificio lontano per il secondo. È facile comprendere come il riconoscimento visuale diretto, senza l'uso di markers appositi, sia immune dalle problematiche di cui essi soffrono e pertanto rappresenti un'alternativa migliore che sta andando incontro ad un rapido sviluppo e, come vedremo più avanti, ad una sempre maggiore diffusione.

How it works: the technical details

Nel caso di VISITO, dato il target dell'applicativo, anziché usare il termine generico di image recognition, si parla più specificamente di landmark recognition (riconoscimento di luoghi d'interesse storico-artistico). Per approcciare questo problema si fa tipicamente uso di tecnologie di machine learning (apprendimento automatico), la più semplice delle quali è l'algoritmo k-NN (k nearest neighbor). Esso classifica un elemento in base ai k elementi più vicini che appartengono in maggior numero ad una categoria, partendo dal presupposto che esista una serie predefinita di categorie tra cui scegliere.

Questa categorizzazione per VISITO ha significato costruire degli insiemi di immagini per ogni monumento, statua, edificio che si è voluto rendere disponibile agli utenti; le immagini sono state in parte selezionate dal famoso sito di fotografia Flickr, in parte fornite dal partner di sviluppo Alinari, per un totale di circa 200000 elementi.

La tecnologia di ricerca all'interno di database di immagini, simile a quella svolta dalla ricerca immagini di Google, è detta *Multimedia Information Retrieval (MMIR*). Essa si differenzia dalla comune ricerca testuale in quanto la chiave di ricerca, o *query*, non è composta da semplice testo bensì da metadati che possono descrivere audio o immagini. La *MMIR* ha visto un notevole sviluppo a partire dal 2000, anno in cui si è assistito all'emergere di nuove tecniche per estrarre features dalle

immagini, e di conseguenza per creare definizioni matematiche delle stesse, dette anche *image signatures* (firme visuali) (Ritendra Datta, 2008).

Nel caso di VISITO, la nostra query sarà composta dai metadati relativi alla foto che abbiamo appena scattato e che vengono estratti da essa.

La comparazione tra due elementi (nel caso di VISITO due immagini) infatti non avviene direttamente, ma si preferisce utilizzare rappresentazioni matematiche di basso livello come istogrammi, contorni, punti d'interesse etc. In particolare la migliore performance di classificazione si riscontra con le cosiddette local features (caratteristiche locali) utilizzate dall'algoritmo SIFT (Scale-invariant feature transform). Il vantaggio del SIFT (Sinha, 2010) è che è in grado di comparare le features di due immagini indipendentemente dall'angolazione, scala, rotazione e illuminazione con cui il soggetto appare in esse: questo significa che, per esempio, due foto del Duomo di Pisa scattate a diversa distanza e posizione, in diverse condizioni di luce, possono comunque venire correttamente classificate e associate alla scheda descrittiva.











Figura 5: SIFT e indipendenza tra orientamento delle immagini

Significa in definitiva che, se la foto scattata dal turista contiene un landmark presente nei database di VISITO, esso viene correttamente identificato.

Nel paragrafo successivo citiamo tra le tecnologie analoghe il framework OpenCV, vale la pena sottolineare che la filosofia per le *query* è opposta a quella di VISITO: infatti vengono cercate le feature di ogni elemento del database all'interno della foto scattata, mentre VISITO esegue la ricerca contraria andando a individuare le feature della foto scattata negli elementi del proprio database.

Who else?

Un prototipo di applicazione per il riconoscimento di landmark era stato sviluppato su piattaforma Symbian e Java nel 2008 dal Computer Vision Laboratory all'interno dell'ETH, l'Istituto Federale Svizzero per la Tecnologia (Till Quack, 2008). All'epoca si presentavano diverse difficoltà di ordine pratico: il GPS integrato era ancora al di là dell'orizzonte, quindi venne implementato un sistema basato sulla triangolazione delle celle individuate dai ripetitori del segnale telefonico. La navigazione su Internet era supportata tramite la tecnologia WAP, e si consentì anche di far inviare la foto scattata tramite MMS; venne anche fatto un tentativo di estrazione delle *feature* dell'immagine in locale, con un algoritmo non ottimizzato, ma i tempi si rivelarono troppo lunghi (circa 10 secondi). L'estrazione delle *feature* avveniva quindi su server remoto, utilizzando una versione estesa delle SIFT, detta SURF, e il risultato finale era un link alla pagina di Wikipedia corrispondente al landmark individuato.

Pur con tutti i limiti del caso, l'esperimento è stato uno dei primissimi tentativi a portare discreti risultati nonostante le risorse – limitate – dell'epoca.



Figura 6: demo del softwarre Symbian sviluppato da ETH su un Nokia N70

Nella realtà odierna di tutti i giorni abbiamo già esempi dell'utilizzo commerciale dell'image recognition, spesso associata alla realtà aumentata: Metaio, una software house tedesca nata nel 2003, ha recentemente rilasciato Metaio Creator, un meta-software in grado di sviluppare applicativi che sfruttano la realtà aumentata, in pochi minuti. Qualsiasi oggetto può essere reso riconoscibile e sovraimpresso con immagini tridimensionali attraverso la camera di uno smartphone. Nel 2011 la famosa catena di negozi Macy's – utilizzando proprio Metaio – ha avviato una campagna promozionale natalizia molto innovativa: i clienti potevano farsi fotografare assieme alla versione virtuale delle sue mascotte (Kats, 2011).



Figura 7: L'app Believe-O-Magic di Macy's

Due anni dopo Ikea ha sviluppato un'app per smartphone in grado di riconoscere le pagine del proprio catalogo, e mostrare i prodotti in 3D sul display; è stata l'app commerciale più scaricata dell'anno, pur venendo lanciata a Luglio (Anonimo, Il nuovo catalogo in realtà aumentata dell'Ikea, 2012). Il successo è stato tale che la multinazionale svedese ha deciso di replicare l'iniziativa con il catalogo 2014 (Ponzio, 2013), ma si tratta anche di una strategia commerciale per risparmiare spese: gli articoli del catalogo sono infatti virtuali, disegnati al computer per ridurre i costi e soprattutto ridurre la possibilità di resi dei clienti a causa di articoli ordinati erroneamente.

La catena di negozi Mercatone Uno si è invece affidata al software Aurasma per il proprio catalogo in realtà aumentata: inquadrando alcune pagine partono video, gallerie di immagini, ecc (Russo, 2012).

Un'alternativa più generalista a Metaio è il framework OpenCV, disponibile su Android e iOS, che fornisce un insieme di librerie per la programmazione; non è specificamente orientato al riconoscimento visuale né alla realtà aumentata (serve genericamente a lavorare in ambito di computer vision), ma si può utilizzare per facilitare il lavoro dei programmatori in tal senso.

Su piattaforma Android inoltre viene fornito OpenCV Manager, un'app contenente le librerie OpenCV precompilate: in tal modo altre app possono farne direttamente uso senza doverle integrare al proprio interno.

L'ecosistema Apple è già ricco di diverse app molto interessanti per iOS, tra tutte sicuramente spicca CamFind (Broida, 2013), che si propone come diretto concorrente di Google Goggles. Tra le caratteristiche più interessanti:

- Riconoscere anche oggetti generici, per esempio una tazza di caffè, e offrire traduzioni in diverse lingue
- Georeferenziare l'immagine e suggerire punti d'interesse rilevanti nelle vicinanze: un altro ristorante, i negozi dove acquistare l'oggetto fotografato, etc.
- Ricerca Google di risultati rilevanti automatizzata
- Confronto prezzi

Si fa notare anche l'app ArtFinder (Barnett, 2011), più vicina all'ambito artistico-culturale di VISITO, e dedicata al riconoscimento delle opere pittoriche; sviluppata dalla britannica Dulwich Picture Gallery, consente di scattare foto ai dipinti esposti nella galleria e visualizzare sul proprio smartphone tutte le informazioni sull'opera e sull'autore. Non manca la possibilità di condividere e commentare i propri risultati attraverso i social network. ArtFinder è stata sviluppata dall'ex direttore operativo del famoso sito Last.fm, Spencer Hyman, e finanziata dal fondatore di Linked.in Reid Hoffman: si prevede un futuro di espansione, con il supporto alle opere esposte in altre gallerie d'arte londinesi e più in generale britanniche. Già adesso sono state catalogate ben 500.000 opere d'arte, un numero impressionante.

L'impegno di Google nel settore della ricerca visuale, iniziato con Goggles e proseguito con l'offerta della ricerca su web per immagini, si è ampliato al social network made in Mountain View, il famoso Google+. Attualmente infatti, gli iscritti possono cercare tra le proprie immagini mediante l'uso di parole chiave, come "gioielleria", oppure "cibo", etc. (Russakovskii, 2013). Sotto al cofano un algoritmo confronta delle immagini di riferimento con quelle dell'utente, alla ricerca di "features" in comune, e infine restituisce i risultati: Google ha preferito giocare prudente e fornire sempre risultati giusti all'utente, anche se a volte incompleti. Lo sforzo tecnologico dietro a questa operazione è davvero notevole, così come la naturalezza della possibilità che ci viene offerta; sembra del tutto naturale, una volta sperimentato, cercare ad esempio tutte le nostre foto di tramonti semplicemente scrivendo la parola: al punto che viene da chiedersi perché non sia sempre stato possibile.

The Dark Side of Image Recognition

La naturalezza dell'interazione è uno dei punti di forza della VR, e più questa diventerà accurata maggiore sarà la sua diffusione; ma c'è anche un rovescio della medaglia cui diversi ricercatori hanno già rivolto la propria attenzione: si tratta dell'annoso problema della privacy.

Già in tempi meno recenti, quando si è cominciato per le prime volte a sentire il termine "realtà aumentata", alcune menti fantasiose hanno immaginato un futuro dove saremo costantemente sommersi da spam virtuale ovunque si poggino i nostri occhi², oppure dove gli hacker ruberanno le nostre informazioni sensibili attraverso una rete pervasiva di realtà aumentata³. Ma non occorre spingersi tanto lontano per avere testimonianze più concrete delle reazioni che queste nuove tecnologie possono provocare.

È del Luglio 2013 la notizia di un caso al limite tra il curioso e l'inquietante. Steve Mann, l'inventore degli occhiali bionici che avrebbero ispirato proprio i Google Glass di Google, è stato assalito mentre si trovava in un McDonalds di Parigi assieme alla propria famiglia (Anonimo, Steve Mann, giallo al McDonald's Danneggiati gli occhiali bionici Eyetap, 2012). Il motivo? Mann porta su di sé EyeTap, un dispositivo per la realtà aumentata in grado di proiettare informazioni aggiuntive direttamente nell'occhio, ma anche di registrare video.

Sarebbe stato questo ad insospettire gli agenti di sicurezza del McDonalds, dando origine a un diverbio in cui avrebbero tentato di rimuovere gli occhiali del professore canadese; tuttavia essi sono fissati al cranio in modo permanente, e sarebbero stati danneggiati dagli uomini della security. È evidente da questo episodio che l'"occhio bionico" è visto ancora con sospetto, complice sicuramente la scarsa informazione che circonda questo tipo di tecnologia.

Nonostante la reazione della sicurezza del locale, conforta tuttavia sapere che numerosi post di accusa e critica siano apparsi sulla pagina Facebook parigina della catena di ristorazione: la diffidenza non è generalizzata insomma, pur se prevalente.

Le poche conoscenze della maggioranza riguardo all'AR (Augmented Reality) inducono alla diffidenza, basti pensare che Google ha dovuto disabilitare il riconoscimento facciale per qualsiasi app sviluppata per i suoi occhiali, i Google Glass. Evidentemente è forte il timore di venire riconosciuti e "taggati" da altre persone, a dispetto della propria privacy (Oremus, 2013).

Diversi locali pubblici inoltre hanno vietato espressamente di indossare i Glass ai propri clienti, scatenando accesi dibattiti su forum e social network, nonché numerosissimi articoli sulle riviste online del settore tecnologico. Caso notevole è stato quello dei casinò americani: chi indossa i Google Glass non può mettere piede nei casinò degli Stati Uniti, da Las Vegas ad Atlantic City, passando per Cincinnati e Cleveland (Oliva, 2013). Prima ancora di sapere se queste funzioni possano o meno essere utilizzate per barare, i responsabili dei casinò americani hanno preferito mettere le mani avanti: quello che i gestori delle case da gioco temono è di non riuscire a dimostrare eventuali imbrogli. Al divieto nelle sale da gioco è seguito quello per gli azionisti di Google, secondo quanto riporta la rivista londinese The Independent (Vincent, 2013); parrebbe che gli stessi inventori della tecnologia temano degli utilizzi impropri. Google tuttavia avrebbe smentito la notizia, precisando che il divieto si estende all'utilizzo di tutte le apparecchiature elettroniche. Si può comunque prevedere che il tema tornerà alla ribalta ad ogni nuovo passo avanti nello sviluppo di queste tecnologie, che sembrano stimolare sì grande interesse, ma anche un certo timore.

² Augmented (hyper)Reality: Domestic Robocop

³ SIGHT

Conclusion: What is the best, what are the differences?

Il parco applicativo si presenta piuttosto variegato, con sfaccettature sia commerciali che informativo-culturali, e viene spontaneo chiedersi chi offre l'app migliore. Un confronto diretto è impossibile, al momento sembra che VISITO sia l'unico a fornire Visual Recognition specificamente nell'ambito della Cultural Heritage, con una copertura di opere d'arte non limitata a una singola città, esposizione, o genere artistico. ArtFinder è dedicato ad una sola galleria d'arte e, cosa forse più importante, alla sola piattaforma Apple. Google Goggles soffre della propria impostazione generalista vero il riconoscimento, così come CamFind, cosa che impedisce una catalogazione più completa in ambito artistico. Si può notare questa differenza specialmente fotografando una statua: VISITO riconosce e dà informazioni sull'opera d'arte, Goggles si limita a suggerire il luogo d'interesse in cui essa è collocata.

D'altra parte le numerosissime app di guida virtuale alle città italiane ed estere, pur competendo con VISITO quanto alla qualità e quantità di utili informazioni, mancano completamente della funzione di VR, e sono in questo senso inferiori.

Sembra di poter concludere che VISITO è superiore in ogni ambito, grazie alla propria specializzazione, ma è meglio fare un'analisi più approfondita. Pur offrendo un parco opzioni integrato e mirato, presenta molti aspetti migliorabili, nei quali potrebbe prendere spunto dalla concorrenza.

Vediamone alcuni:

Uso del GPS:

VISITO localizza l'utente sulla mappa di Google e la correda dei relativi punti d'interesse, ma non offre un servizio importantissimo, ovvero la navigazione. Eppure basterebbe appoggiarsi al navigatore integrato offerto sia da Android che iOS. Moltissime app turistiche hanno questa utile funzionalità;

Image Recognition:

Sicuramente da perfezionare l'interfaccia, scattare foto a distanza ravvicinata richiede almeno un paio di focus manuali successivi. Altre piccole migliorie ai menu, pur secondarie, migliorerebbero ancor di più l'esperienza utente;

Galleria di foto:

Le foto a schermo intero non sono integrate in una galleria, forzando l'utente a una navigazione tra l'elenco foto e gli ingrandimenti. Questa funzionalità è comunissima e andrebbe davvero implementata;

• Il sito desktop:

Il sito desktop che fa da companion all'app potrebbe tranquillamente essere integrato con l'applicazione stessa, grazie alla potenza di calcolo degli smartphone attuali. Del resto molto spesso le visite si pianificano nel dettaglio una volta raggiunte le località d'interesse.

• Risultati della ricerca:

Offrire soltanto foto e descrizione testuale, al momento, è un'offerta povera rispetto alle possibilità tecnologiche. Video, 3D, supporto vocale, link a risorse esterne (es. Google) possono e dovrebbero essere aggiunti per arricchire l'esperienza utente.

• Obbligo di connessione:

La connessione ai server di VISITO è indispensabile per il riconoscimento delle foto, ma non per tutte le altre funzioni dell'app come la consultazione delle schede o della mappa (per quest'ultima basta il GPS). Sarebbe utile mantenere le schede e la mappa in locale, per un risparmio sia del consumo dati che della batteria (la connessione 3G tassa molto gli smartphone, e LTE è ancora poco diffuso), come fanno molte app dedicate al turismo.

Questi i principali fattori che rendono VISITO un'applicazione molto interessante, ma sicuramente aperta a numerosi miglioramenti. Le motivazioni si trovano nella precoce nascita di VISITO, creata quando ancora gli smartphone avevano tastiere fisiche, poca potenza di calcolo e schermi di dimensioni ridotte, per cui implementare i suggerimenti di cui sopra sarebbe stato alternativamente impossibile, poco performante, o avrebbe prodotto un'esperienza utente poco piacevole.

I tempi per un'evoluzione di VISITO sono maturi (basti pensare alla diffusione dei tablet), purtroppo i partner principali che hanno contribuito alla nascita del progetto non sono più in grado di portarlo avanti, per questo sarebbe bene che si trovassero nuove energie, al fine dare linfa vitale al team di sviluppo e consentire a VISITO l'evoluzione che merita. Gli sviluppatori, interrogati sul futuro della loro creatura durante una conferenza all'Università di Pisa, si sono mostrati ben consapevoli dei sui limiti attuali e hanno loro stessi spiegato le cause; circola un cauto ottimismo sulla possibilità di nuovi player che potrebbero far ripartire i lavori.

Bibliografia

- Anonimo. (2012, Agosto 24). *Il nuovo catalogo in realtà aumentata dell'Ikea*. Tratto da Realtà aumentata: http://goo.gl/fFCiMu
- Anonimo. (2012, Luglio 17). *Steve Mann, giallo al McDonald's Danneggiati gli occhiali bionici Eyetap*. Tratto da LaRepubblica.it: http://goo.gl/zKLH18
- Barnett, E. (2011, Luglio 22). *Artfinder: new app recognises paintings*. Tratto da The Telegraph: http://goo.gl/OkwuTu
- Broida, R. (2013, Giugno 3). *CamFind turns your iPhone camera into a search engine*. Tratto da cNet: http://goo.gl/m3hCG
- Colleen, H. (2010, Ottobre 20). *Just Picture It: Mobile Visual Recognition Software*. Tratto da New Media Resource Studio FA2: http://goo.gl/2IKuh
- David Marimon, T. A. (2010). *PEREY Research and Consulting*. Tratto da Mobile Visual Recognition, the future of Mobile AR: http://goo.gl/UtYjpg
- Dunay, P. (2013, Aprile 4). *Using Visual Recognition to Tap into the Consumer Mindset*. Tratto da Socialmedia today: http://goo.gl/0qZhS
- Kats, R. (2011, Novembre 7). *Macy's strengthens holiday mobile efforts via augmented reality*. Tratto da Mobile Marketer: http://goo.gl/rl57q
- Oliva, M. (2013, Giugno 7). *Google Glass, vietato l'utilizzo nei casinò*. Tratto da GEEKiSSIMO: http://goo.gl/S2uSMq
- Oremus, W. (2013, Giugno 3). "Don't Be Creepy": Google Glass Won't Allow Face Recognition. Tratto da Slate: http://goo.gl/jCzjfr
- Ponzio, S. (2013, Agosto 9). *Ikea e arredi casa con smartphone e tablet*. Tratto da Focus.it: http://goo.gl/DG9904
- Ritendra Datta, D. J. (2008). Image Retrieval: Ideas, Influences, and Trends of the New Age. *ACM Computing Surveys*, 3-5;14-18.
- Russakovskii, A. (2013, Maggio 21). Google+'s Best New Unadvertised Feature: Photo Search With Visual Recognition Try It On Your Own Pictures And Be Amazed. Tratto da ANDROID POLICE: http://goo.gl/dLfOd
- Russo, D. (2012, Ottobre 6). *Mercatone Uno Presenta il Catalogo Autunno con Realtà Aumentata*. Tratto da MobilPro Arredamento e Pubblicità: http://goo.gl/XaU5uR
- Sam Habibi Minelli, G. A. (2011, Dicembre). VISITO Tuscany: uno strumento di valorizzazione e marketing territoriale. Tratto da VISITO Tuscany: http://goo.gl/JZ92hE
- Sinha, U. (2010, Maggio 14). SIFT: Scale Invariant Feature Transform. Tratto da AI Shack: http://goo.gl/CK6NdG
- Till Quack, H. B. (2008). *Object Recognition for the Internet of Things*. Tratto da ETHZ Computer Visual Laboratory: http://goo.gl/8ZKvab
- Vincent, J. (2013, 7 Giugno). *Google bans Glass at shareholder meeting*. Tratto da The Independent: http://goo.gl/DN09I