

# Progetto: Fruizione di contenuti attraverso immagini panoramiche

## Rapporto tecnico: Studio della transizione tra immagini nel passaggio tra panoramiche adiacenti nella sequenza

# Navigazione tra immagini panoramiche

Data una sequenza di immagini panoramiche acquisite lungo un percorso approssimato da una curva C(t)=(x(t),y(t)) su un piano XY si vuole progettare un meccanismo di transizione tra panoramiche che consenta di muoversi tra panoramiche adiacenti nella sequenza riducendo la discontinuità percepita dall'osservatore nella fase di transizione.

Sia P1, ..., Pn la sequenza di immagini panoramiche acquisite in corrispondenza dei punti C(t1), ..., C(tn) sul percorso, con t1 < t2 < ... < tn. Si assume che per ciascuna coppia di panoramiche adiacenti (Pk, Pk+1) siano state individuate delle direzioni di navigazione che individuano, in coordinate immagine, direzioni omologhe sulle due panoramiche.

Figura 1

Per la generica coppia di panoramiche (Pk, Pk+1) tali direzioni di navigazione sono quindi esprimibili come una quadrupla Q(Pk, Pk+1) = (xpk, xak, xpk+1, xak+1) che individua:

* xpk : l'ascissa della direzione di navigazione nel partire dalla panoramica Pk per spostarsi su Pk+1 ;
* xak : l'ascissa della direzione di navigazione nell'arrivare alla panoramica Pk provenendo da Pk+1 ;
* xpk+1 : l'ascissa della direzione di navigazione nel partire dalla panoramica Pk+1 per spostarsi su Pk ;
* xak+1 : l'ascissa della direzione di navigazione nell'arrivare alla panoramica Pk+1 provenendo da Pk .

Da notare che questa ipotesi prevede che per la generica panoramica Pk vengano individuate quattro direzioni di navigazione: una di partenza per la panoramica precedente Pk-1, una di arrivo dalla panoramica precedente Pk-1, una di partenza per la panoramica successiva Pk+1 ed una di arrivo dalla panoramica successiva Pk+1. La Figura 1 evidenzia i valori della quadrupla Q(P1, P2). Sulla stessa figura sono anche evidenziate in giallo alcune regioni rettangolari che approssimano la scena visualizzata, ossia la porzione di panoramica visualizzata nel pannello dell'applicazione di navigazione. La dimensione della scena visualizzata dipende da due fattori: i) la dimensione della finestra all'interno della quale è in esecuzione l'applicazione di navigazione e ii) il fattore di zoom scelto dall'utente per visualizzare il contenuto della panoramica.

La metafora adottata per supportare la navigazione tra panoramiche si basa sul monitoraggio dell'azione di zoom effettuata dall'utente. Se questa azione è messa in atto quando la scena visualizzata è lontana dalle direzioni di partenza della panoramica corrente (partenza verso la panoramica precedente o quella successiva) allora l'azione di zoom viene tradotta in un effettivo ingrandimento del contenuto della scena visualizzata e di conseguenza la regione rettangolare che la approssima sulla panoramica risulterà avere una estensione minore rispetto alla visualizzazione con zoom 1:1. Se invece l'azione di zoom è messa in atto dall'utente quando la scena visualizzata è centrata su una delle direzioni di partenza della panoramica corrente, per esempio verso la panoramica successiva, allora l'azione di zoom innesca la transizione dalla panoramica corrente a quella successiva.

Si vuole quindi definire un modello per il processo di transizione che consenta di ridurre la discontinuità percepita dall'utente nella fase di aggiornamento del contenuto della scena visualizzata passando dalla panoramica corrente alla successiva. Uno scenario di esempio è mostrato nella Figura 2.

Figura 2

A sinistra è mostrato il contenuto della scena visualizzata nella panoramica corrente (zoom 1:1), in corrispondenza della direzione di partenza verso la panoramica successiva il cui contenuto nella direzione di arrivo (zoom 1:1) è mostrato a destra. Risulta evidente che una transizione tra le due scene che prevedesse la sostituzione dell'immagine di sinistra con quella di destra nel passaggio dalla panoramica corrente alla successiva causerebbe una forte percezione di discontinuità nell'osservatore. In particolare, tale discontinuità è causata dal fatto che la panoramica successiva rappresenta i contenuti di scena relativi alla direzione di navigazione ad un livello di ingrandimento maggiore rispetto alla panoramica corrente, in quanto il punto di scatto della panoramica successiva è più vicino agli oggetti della scena rispetto al punto di scatto della panoramica corrente. Questo suggerisce che una riduzione della discontinuità nella transizione tra le due scene possa essere ottenuta effettuando un'azione di zoom sulla panoramica corrente sino a visualizzare una scena in cui le stesse strutture presenti nella scena della panoramica successiva sono rappresentate alla stessa dimensione.

Figura 3

In Figura 3 l'immagine di sinistra mostra la scena visualizzata nella panoramica corrente (zoom 1:1) con evidenziata la regione rettangolare che meglio approssima il contenuto della scena visualizzata nella panoramica successiva (Figura 2 a destra). Il contenuto della zona rettangolare corrisponde alla scena visualizzata nell'immagine corrente se si applica un fattore di zoom 1:2 nella direzione di navigazione. In questo caso, il contenuto della scena visualizzata nella panoramica corrente (zoom 1:2) è visualizzato nella Figura 3 a destra. Risulta evidente come l'effetto di transizione tra le immagini a destra nelle Figure 2 e 3 producano una percezione di discontinuità nettamente inferiore a quella associata alla transizione tra le due immagini di Figura 2.

Il modello di transizione dalla generica panoramica Pk alla panoramica Pk+1 si articola quindi nei seguenti passi:

* **Sub-scene matching**. Il contenuto della scena della panoramica Pk+1 in corrispondenza di xak+1 viene analizzato da un modulo automatico in modo da trovare in corrispondenza di xpk nella panoramica Pk la porzione di scena più simile al contenuto della scena della panoramica Pk+1. In particolare, la dimensione della porzione di scena individuata determinerà il fattore di zoom Zk da applicare nella visualizzazione della panoramica Pk prima di attivare la transizione alla panoramica successiva Pk+1.
* **Scene visualization and transition**. In fase di navigazione del contenuto della panoramica Pk una azione di zoom nella direzione di transizione porterà ad un effettivo ingrandimento del contenuto della scena corrente solo sino al raggiungimento del fattore di zoom Zk raggiunto il quale il modulo di navigazione presenterà il contenuto della scena della panoramica Pk+1.

Da notare che il passo di Sub-scene matching viene eseguito una volta per tutte in una fase di analisi della sequenza di panoramiche che è svolta a monte della fase di navigazione la quale sfrutta, per implementare la Scene transition, i parametri individuati nel passo di Sub-scene matching. Nel seguito vengono descritti i due processi di Sub-scene matching e Scene transition.

## Sub-scene matching

Date due immagini Ik ed Ik+1. si vuole trovare la sottoregione rettangolare di Ik il cui contenuto è più simile ad Ik+1. Supponiamo di adottare la seguente notazione:

* wk+1 ed hk+1 siano la larghezza ed altezza in pixel dell'immagine Ik+1.
* SubImg(I, x, Δx, y, Δy) sia la funzione che restituisce la porzione rettangolare di immagine individuata dal vertice in basso a sinistra di coordinate (x,y) e dal vertice in alto a destra di coordinate (x+Δx, y+Δy)
* D(I1 , I2) sia la funzione distanza tra immagini che restituisce zero nel caso che il contenuto delle due immagini sia lo stesso ed altrimenti un numero maggiore di zero. Senza perdita di generalità si può assumere che qualora le due immagini non abbiano la stessa dimensione una delle due venga ridimensionata in modo da avere la stessa dimensione dell'altra

Figura 4

La sottoregione rettangolare è identificata attraverso la minimizzazione del seguente funzionale di costo rispetto alle tre variabili (x,y,η):



dove i valori (xo,yo,ηo) delle variabili che minimizzano il funzionale di costo rappresentano rispettivamente le coordinate del vertice in basso a sinistra della sottoregione identificata (xo,yo) ed il fattore di zoom (Z=1/ηo) da applicare alla sottoregione per ridimensionarla alla grandezza di Ik+1.

Il valore della funzione D(I1 , I2) è calcolato estraendo gli istogrammi delle due immagini e calcolando il valore della distanza quadratica media tra bin corrispondenti. Nella sperimentazione effettuata sono stati considerati istogrammi estratti dalla componente di luminosità, di cromaticità e saturazione. Tra questi quelli con maggior efficacia sono risultati gli istogrammi estratti dalla componente di luminosità. Per ridurre i tempi di calcolo associati alla minimizzazione del funzionale di costo è stata adottata la tecnica degli Istogrammi Integrali (F. Porikli. Integral histogram: a fast way to extract histograms in Cartesian spaces. In Proc. of IEEE Int. Conf. on Computer Vision and Pattern Recognition (CVPR), 2005).

Nella seguente figura sono mostrati alcuni risultati del processo di selezione automatica della sottoregione. Su ogni riga sono riportate tre immagini; quella a sinistra rappresenta la scena della panoramica corrente, quella al centro la sottoregione della panoramica corrente che meglio approssima la scena della panoramica successiva che infine è rappresentata nell'immagine più a destra. In corrispondenza dell'immagine centrale è anche mostrato il fattore di zoom da applicare alla panoramica corrente per produrre la visualizzazione della sottoregione individuata.

Figura 5

I risultati mettono in evidenza come il modello descritto sia in grado di individuare efficacemente la sottoregione della scena corrente che meglio approssima la scena della panoramica successiva nel caso in cui non intervengano significative occlusioni tali da nascondere nella panoramica corrente parti rilevanti di scena visibili nella panoramica successiva. Questo è quanto invece si verifica nella prima serie di immagini mostrate in Figura 5: a causa della prima macchina parcheggiata le macchine osservabili nella panoramica successiva ed anche lo spazio vuoto di fronte alla tettoia sono del tutto occlusi nella panoramica corrente. In questo caso la sottoregione individuata nella scena corrente è solo parzialmente corrispondente alla scena mostrata nella panoramica successiva.

## Scene visualization and transition

La realizzazione dell’interfaccia utente che consente la navigazione dell’ambiente per mezzo di transizioni fra foto panoramiche è stata implementata usando la libreria 3D ‘opensource’ [three.js](http://threejs.org/) basata su HTML5, CSS3, javascript e WebGL. La resa tridimensionale di ogni scena viene simulata mappando la foto sferica all’interno di una sfera e ponendo la telecamera, e quindi il punto di vista,al centro della stessa. La foto sferica diventa dunque la texture delle maglie che costituiscono la geometria dell’oggetto tridimensionale.

Il browsing di ciascuna scena è ottenuto attraverso due azioni principali: il ‘drag’ dell’immagine consente la modifica dello scorcio di panorama visibile in quel momento, mentre l’uso della ‘mouse wheel’ permette lo zoom all’interno dello scorcio. La posizione della scena viene gestita attraverso le coordinate di latitudine e longitudine, rese coerenti con la direzione della camera, mentre il fov (field of view) evidenzia in ogni momento la quantità di zoom presente. Inoltre per evitare problemi come in fig. 2 sono stati aggiunti nel programma alcuni ‘constraints’ di visualizzazione.



Figura 6

Alcuni constraints consentono per esempio di nascondere il buco nella panoramica causato dal punto da cui è stata scattata

La navigazione fra una foto sferica e le successive è stata invece realizzata attraverso la sostituzione dinamica della panoramica utilizzata come texture all’interno della sfera, secondo i parametri di interazione utente. Allo stato attuale l’aggiornamento della texture avviene gestendo attraverso delle soglie il livello di zoom ottimale a cui effettuare la transizione. Questa può avvenire se nella direzione dell’obbiettivo della camera sia prevista una foto corrispondente con cui sostituire quella attuale, altrimenti il ‘field of view’ non viene ulteriormente diminuito (Fig. 7 e Fig .8)



Figura 7

pano1 al massimo zoom in direzione di 2 e pano2 al caricamento da 1



Figura 8

pano3 al massimo zoom in direzione di 2 e pano2 al caricamento da 3

Per ogni transizione viene controllato inoltre se la latitudine e longitudine attuali rientrano in qualunque dei possibili intervalli in cui sia stata prevista una nuova texture, in tal caso questa viene sostituita alla precedente ripristinando la corretta prospettiva.