

Pràctiques de VC/PSIV

Felipe Lumbrieras Ruiz

Lab 3: Mosaics

3.1 Introducció

Aquesta pràctica està basada en la pràctica “Mosaics panoràmics” redactada per en Joan Serrat a l’assignatura de Processament d’Imatges de l’Enginyeria Informàtica de la UAB.

Avui dia les càmeres fotogràfiques amb pel·lícula de negatiu han passat a la història, i la fotografia és digital. Les càmeres actuals no només emulen els controls de diafragma i temps d’exposició de les antigues càmeres, sinó que, en lliurar-nos imatges digitals, permeten un postprocès a l’ordinador (o a la càmera mateixa) molt més versàtil i fàcil d’aplicar que no pas la manipulació de negatius. El camp de la **fotografia computacional**, a cavall de la visió per computador i el processament d’imatges, s’ocupa de l’estudi d’aquests tipus de processos.

Una de les moltes possibilitats que obre el processament digital d’imatges fotogràfiques és la construcció de *mosaics panoràmics*. Són imatges construïdes combinant una sèrie d’imatges originals, de manera que es cobreixi un camp de vista angular molt més ampli que qualsevol d’elles. Vénen a emular el resultat que s’aconsegueix amb els objectius gran angular. Però a diferència d’ells, ara es poden aconseguir imatges que cobreixin fins a tot el camp de vista angular: 360° verticalment i horitzontalment. La figura 7 mostra un panorama d’una cruïlla de carrers: el camp de vista angular horitzontal és de 360°, mentre que el vertical és d’uns 40° o 50°, o sigui, no s’ha ampliat. La figura 8 mostra un panorama semblant, però de l’accés a l’EE.

Podeu trobar moltíssims exemples de mosaics panoràmics a la web. A part del seu interès estètic, es fan servir per oferir visites virtuals a paisatges, ciutats, habitatges, carrers (Google street), etc. Per fer els vostres propis mosaics, proveu la versió de demostració del programa Autostitch de www.autostitch.net o feu servir les apps de càmera panoràmica que porten molts mòbils.

La manera de generar mosaics panoràmics és “ajuntant” adequadament una sèrie d’imatges capturades des d’un mateix punt però rotant la càmera sobre si mateixa, sense traslladar-la. Dit d’una altra manera, si únicament hi ha una rotació 3D de la càmera respecte del seu centre òptic (vegeu les transparències del tema Transformacions geomètriques) llavors és possible deformar cada una d’elles per tal que es pugui “ajuntar” amb les seves dues veïnes, també deformades de la mateixa manera. Ajuntar en el sentit de què un cop deformades les imatges, a les vores tenen la mateixa geometria i per tant podem efectuar una simple translació per encaixar-les. Una altra qüestió, com veurem, és aconseguir que a les junes no es notin canvis

de contrast.

Si només girem la càmera respecte d’un eix vertical, és a dir, perpendicular al terra, la sèrie d’imatges es pot transformar en un panorama com el de les figures 7, 8. Observeu que les línies rectes horizontals en l’escena —paral·leles al terra— apareixen corbades, però les verticals no. Si a més horitzontalment, rotem la càmera verticalment per augmentar el camp de vista angular en aquesta direcció, en fer el panorama també les línies verticals deixen de ser rectes. Per raons que quedaran clares més endavant, aquest és el preu que cal pagar per poder construir panorames cilíndrics.

Aquest problema *sembla* senzill. Després de tot, segur que algun dia heu intentat construir un mosaic de dues o tres fotografies impresa, i només traslladant-les i rotant-les manualment *més o menys* ho haureu aconseguit. No obstant això, la relació matemàtica entre els sistemes de coordenades de les imatges adquirides tal com s’ha explicat no és la composició d’una rotació en el pla i una translació (més escalat si heu canviat la distància focal): si ho proveu, veureu que així no hi ha manera d’ajuntar-les exactament (figura 10).

3.2 Objectius

- Aprendre a realitzar diferents tipus de panorames: plans (gran angular) i cilíndrics (circulars).
- Entendre la geometria implicada en cadascuna de les projeccions
- Proposar un algorisme de *blending* per a tractar les superposicions.
- Treballar diferents maneres de fer la correspondència de punts entre imatges.

3.3 Tasques

Les tasques que s’han de realitzar són:

1. Escollir i capturar les imatges amb què treballarem.
2. Projectar les imatges per tots dos tipus de panorames.
3. Calcular el desplaçament o la transformació necessària que porta una imatge a la seva veïna.
4. Unir les imatges formant una única imatge resultat.
5. Millorar el procediment dels punts que ho permetin.

3.3.1 Escollir i capturar les imatges

Per assolir els objectius es proposa fer panorames amb dos conjunts d’imatges, unes que anomenarem problema i d’altres que heu d’adquirir vosaltres.

Per les imatges problema agafareu un conjunt d’imatges, només un d’entre aquests: cnm, ee, iiia i ciències. Mireu-los tots i agafeu el que més us agradi.



Figura 7: Panorama de 360° .



Figura 8: Panorama “ee”. El camp de vista angular horitzontal és 360° , el cantó dret continua al cantó esquerre.



Figura 9: Imatges originals del panorama “ee”

Per fer la vostra adquisició féu una rotació de la càmera al voltant del centre òptic i enfocant a objectes llunyans si és possible. Mireu com són les seqüències anteriors.

3.3.2 Projectar les imatges

S’han de fer dos tipus de mosaics. El primer tipus de panorama, que anomenarem pla, constarà d’una concatenació de 3 imatges, per fer senzill el problema. Hem d’escollir

un conjunt d’elles que obrin menys de 180° . Les escollirem dels mateixos conjunts d’abans, de les imatges problema i de les vostres. La imatge del mig la deixarem sense transformar i a les imatges de les bores les hi aplicarem una homografia per tal d’ajustar la geometria de la unió. Aquesta homografia es pot treure de la correspondència entre punts.

El segon que anomenarem cilíndric on el conjunt d’imatges serà gran, inclús pot donar la volta sencera. En aquest cas es farà una projecció cilíndrica en totes les imatges de la seqüència. Volem que les imatges planes que tenim quedin projectades a sobre d’un cilindre, com es veu a la figura 11. Busqueu informació de com es fa una projecció cilíndrica (apèndix). Utilitzeu el que heu après en el tema de Transformacions geomètriques per fer les projeccions.

3.3.3 Calcular el desplaçament o transformació

La correspondència entre punts de les imatges la farem de forma manual. Per aquesta part escollirem punts destacats en la primera imatge que podem reconèixer en la



Figura 10: Manca de correspondència geomètrica en rotar i traslladar simplement

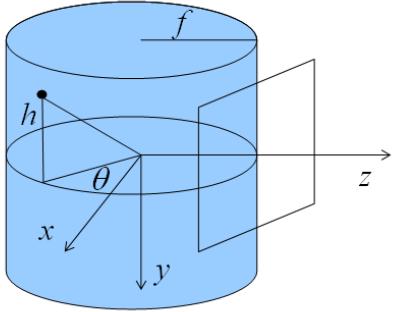


Figura 11: Projecció d'una imatge plana en un cilindre, elements involucrats

segona. Si ja tenim fetes les projeccions de les imatges el que obtindrem amb aquestes correspondències serà una translació horitzontal que farà coincidir les imatges. Si cliquem més d'un parell de punts el que haurem de tenir són valors similars, per tant ens quedarem amb la mitjana un cop hagim tret els valors erronis.

En el cas del panorama pla farem el mateix procediment si ja tenim fetes les projeccions, però el més habitual és fer servir els punts de correspondència per resoldre un sistema i obtenir l'homografia que hem d'aplicar. L'algorisme que es fa servir és el DLT (*Direct Linear Transform*) que donat un seguit de parelles de correspondències calcula l'homografia que porta uns punts als altres.

3.3.4 Generar una imatge resultat

En tots dos casos, ja tenim les imatges aïllades i coneixem com les hem de col·locar. Llavors farem una imatge negra base, tan gran com sigui necessari en funció de les mides de les imatges projectades i dels vectors de desplaçament. Sobre aquesta imatge col·locarem les imatges que formaran el panorama.

Ara toca unir-les. La manera més directa és que una imatge tapi els punts comuns que té amb les altres. Per fer això simplement anirem col·locant sobre la imatge negra base les diferents imatges (actualitzant només els píxels de la imatge que estarem tractant en aquell moment). Amb aquesta rutina taparem els píxels comuns anteriors amb els nous.

3.3.5 Millorar alguns punts del procediment

Podem millorar la correspondència fent servir punts detectats automàticament, en comptes de picar els punts un per un fem servir un algorisme que detecta punts que són rellevants per algun motiu (tenen més informació que la resta dels seus veïns). Tot això es veurà més endavant en el curs com a punts característics. Ara podem aprofitar que MatLab porta ja algun d'aquests detectors implementats (en funció de la versió trobareu SURF, MSER, FREAK, BRISK, FAST, ...) El que farem és llençar l'algorisme en totes dues imatges que volem unir i de les possibles

combinacions ens quedarem en les que siguin plausibles amb la transformació que volem fer. Aquesta última part que combina i troba les unions que s'ajusten a un model s'anomena RANSAC i en podreu trobar versions per a MatLab (en les darreres versions forma part de la funció `estimateGeometricTransform` que podeu fer servir directament). Mireu exemples de com es fa servir la combinació de la detecció de punts i el RANSAC per exemple per fer mosaic a partir de vídeos directament.

La part de la unió final de les imatges es pot millorar amb un procés de *blending*. Aquestes combinacions poden ser molt complexes per tal de minimitzar els errors que es puguin donar en aquesta unió. No obstant podem fer servir una primera versió que consisteixi a fer la mitjana dels píxels implicats. Una manera de fer aquest *blending* és fent servir dues imatges base negres. En una d'elles acumulem (sumem) la contribució de les imatges, el que donarà que en aquells píxels on hi hagi més d'una contribució el valor es doblarà possiblement. Amb la segona imatge negra base acumularem imatges de màscares on tindrem un 1 si tenim imatge i un 0 si no en tenim. Per tant en aquells punts coincidents tindrem un valor major a 1. Aquesta segona imatge és el factor pel qual hem de dividir la primera per restablir els valors dins del rang.

3.4 Puntuació, lliurament, llenguatge

Solucionar el problema de fer panorames dóna 6 punts de la nota, repartits d'aquesta forma:

- panorama pla sobre imatges problema (3 punts)
- panorama pla amb imatges vostres (1 punt)
- panorama cilíndric sobre imatges problema (3 punts)
- panorama cilíndric amb imatges vostres (1 punt)

Presentar una solució de blending per a la superposició d'imatges més elaborada que simplement tapar els píxels coincidents d'una amb els de l'altra serà 1 punt.

Fer servir altres estratègies per fer la correspondència entre els punts més enllà de la manera manual també seran 2 punts.

En aquest lliurament no hi ha apartat específic d'extres, intenteu aconseguir-ho tot. Si heu fet un dels panorames, l'altre ja el tindreu mig fet i l'heu resolt per les imatges problemes hauria de ser gairebé automàtic provar-ho amb les vostres imatges. El blending es pot resoldre fàcilment amb una mitjana dels píxels afectats i per la part de correspondència podeu mirar exemples que porta el mateix MatLab on es resol de manera automàtica la generació de mosaics.

Podeu fer l'entrega en MatLab o Python.

Referències

- ◊ Apunts de classe.

Apèndix: projecció cilíndrica

Un cop aplicada la projecció veurem que les imatges deixen de ser rectangulars per passar a tenir una forma de píndola, arrodonides per la part de dalt i de baix, com es veu a la figura 12.

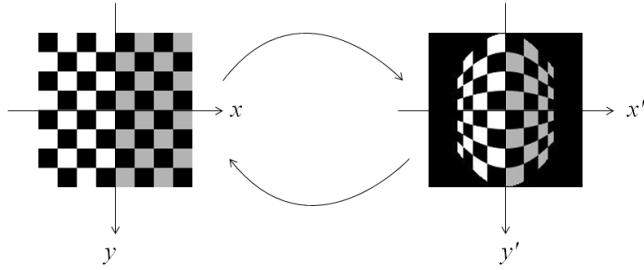


Figura 12: Projecció cilíndrica d'una imatge plana

Les equacions que passen un sistema a l'altre són:

$$\begin{aligned} x' &= s\theta = s \tan^{-1} \frac{x}{f} \\ y' &= sh = s \frac{y}{\sqrt{x^2 + f^2}} \end{aligned}$$

i la transformació inversa:

$$\begin{aligned} x &= f \tan \theta = f \tan \frac{x'}{s} \\ y &= h \sqrt{x^2 + f^2} = \frac{y'}{s} f \sqrt{1 + \tan^2 \frac{x'}{s}} = f \frac{y'}{s} \sec \frac{x'}{s} \end{aligned}$$

Podem agafar el factor d'escalatge s igual a f . Feu proves amb diferents f per obtenir un valor que vagi bé.

Tres coses a tenir en compte. En primer lloc, quina parella d'equacions s'ha d'utilitzar? Mireu els dibuxos de la figura 12 i identifiqueu quin és el vostre input i el vostre output. Feu servir aquelles que hem comentat a teoria que són les adients en aquest tipus de mappings. Encara que per obtenir les mides de l'output necessitarem l'altre conjunt d'equacions aplicades a alguns punts estratègics. En segon lloc, aquestes equacions estan centrades, és a dir que el $(0,0)$ tant de l'espai de partida com el d'arribada estan al mig de la imatge. Les imatges, per contra, tenen l'origen en el píxel esquerre superior, per tant haurreu d'incloure aquestes translacions perquè tot quadri. I per acabar podeu fer servir un codi similar al que teniu a problemes (el de la rotació o el del smile-pou) adaptat convenientment per aquest cas concret.