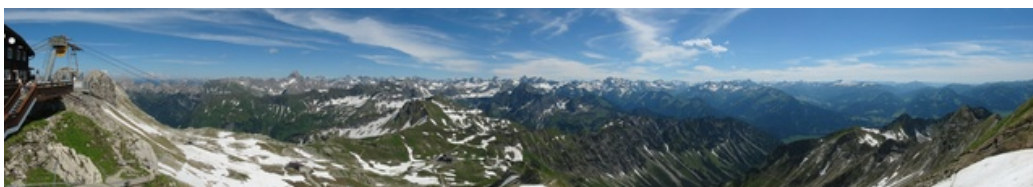


Enganxant fotografies

Gregori Guasp

Segur que més d'una vegada us heu trobat intentant fer una fotografia de la façana d'un edifici i, per culpa de no poder-vos allunyar prou, el camp de la vostra màquina no us ha permès enquadrar-lo tot sencer. És molt possible també que davant d'aquesta situació hàgiu provat de fer un parell, o més, de fotografies parcials per a enganxar-les després i reconstruir així una imatge de l'edifici tot sencer. Sense cap mena de dubte és segur que els resultats que haureu obtingut aniran des d'un resultat simplement mediocre cap a resultats francament desastrosos que no s'assemblen gaire a la imatge dels Alps (obtinguda enganxat 10 fotografies) següent¹



Fixeu-vos que aquest és, exactament, el mateix problema que apareix quan es fan fotografies aèries. S'obtenen un munt de fotografies del terreny mentre l'avió (o el satèl·lit) fa diferents passades i cal acoblar-les totes per a produir un mapa o una imatge de tot el territori. Naturalment, cada una de les fotografies s'haurà obtingut des de posicions i orientacions diferents de forma que sense un tractament posterior és impossible fer coincidir totes les peces d'una forma coherent i precisa.

¹Imatge de Pablo d'Angelo amb llicència *Creative Commons* obtinguda al lloc <http://home.arcor.de/pablo.dangelo/album/>

Posem un exemple

Per a comprovar fins a quin punt es pot solucionar el problema amb una mica de matemàtiques considerarem el parell de fotografies de la figura 1 (una pissarra del despatx d'un conegut).

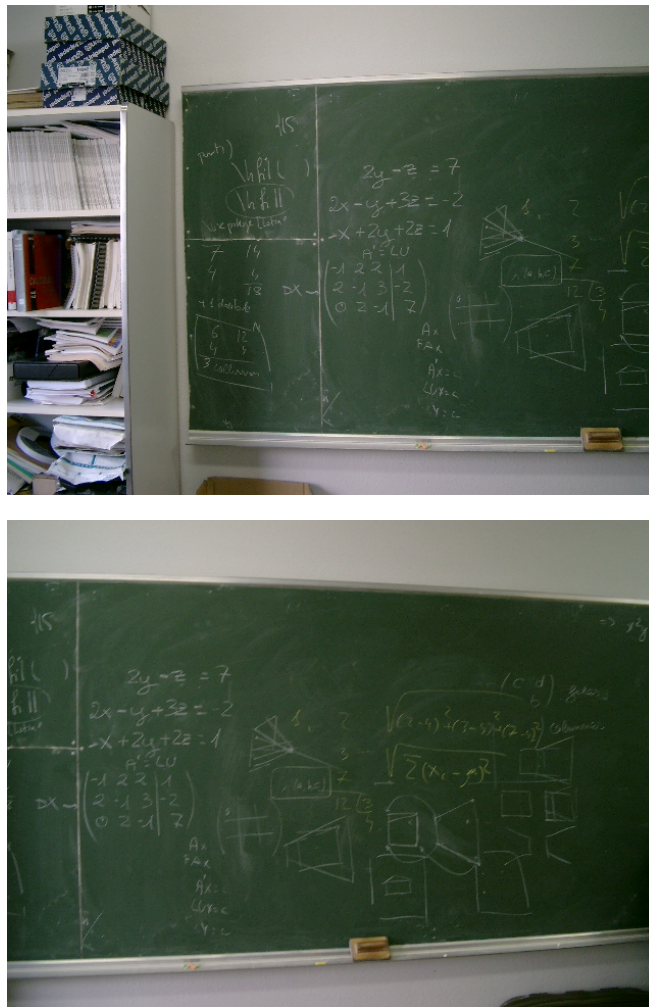


Figura 1:

Simplificarem el problema fixant-nos només en les línies rectes que apareixen i, per a completar una mica els detalls, en el grup de guixos i en

l'esborrador, calcant tots aquests detalls per a obtenir un parell d'esquemes com els de la figura 2.

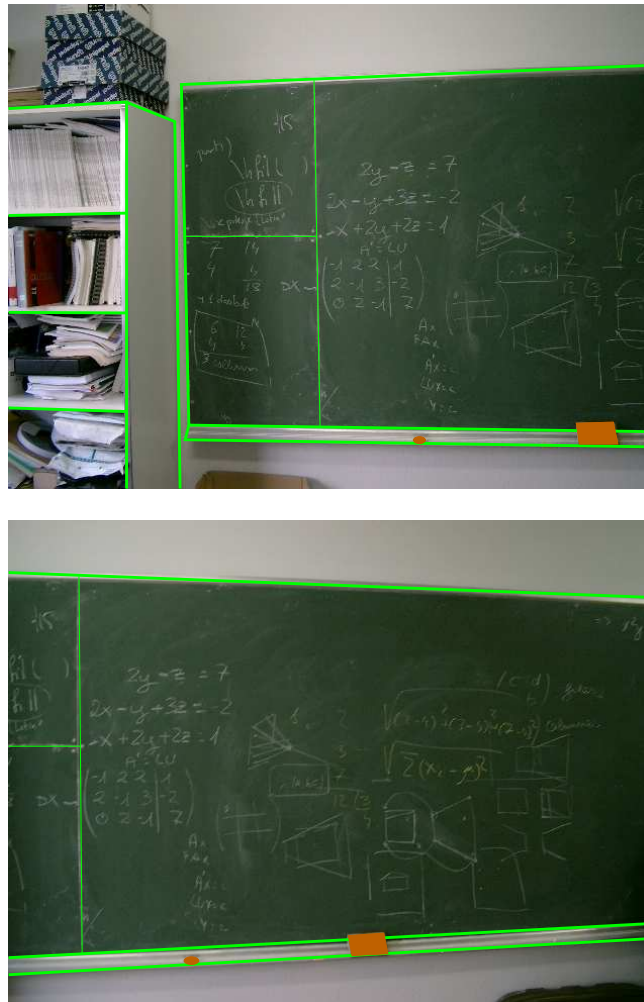


Figura 2:

El resultat final que s'obté després d'aquest procés (figura 3) és una representació prou exacta de les dues fotografies de partida i és d'on en traurem totes les dades per als càlculs posteriors. Noteu que el grup de guixos ha quedat reduït a un únic punt i que de l'esborrador només s'ha marcat l'extrem superior esquerra en els dos esquemes. Observeu també que en l'esquema

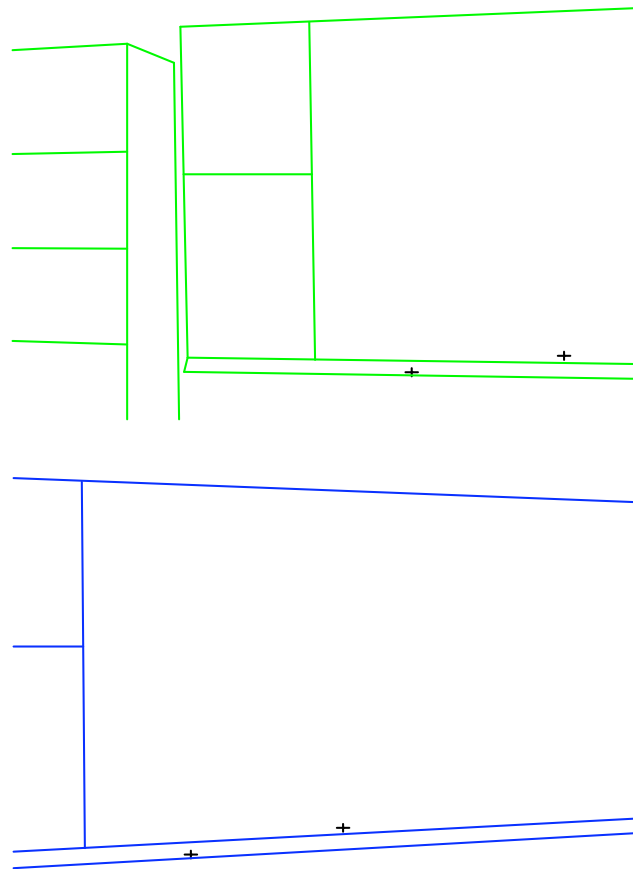


Figura 3:

queda molt clar que mentre que les línies horitzontals dels dos extrems de la pissarra convergeixen cap a l'esquerra en la fotografia superior ho fan cap a la dreta en l'altra i que, per tant, si intentem enganxar les dues fotografies directament una sobre l'altra mai aconseguirem, per exemple, que l'extrem superior o inferior de la pissarra es vegi com una línia recta (fins i tot encara que canviem les ampliacions).

Fixeu-vos finalment que aquesta simplificació permet resumir tota la informació més important de les dues figures en les coordenades dels punts extrems de cada un dels segments que apareixen en els dibuixos i que d'aquesta forma no cal treballar amb tots els *pixels* de les fotografies.

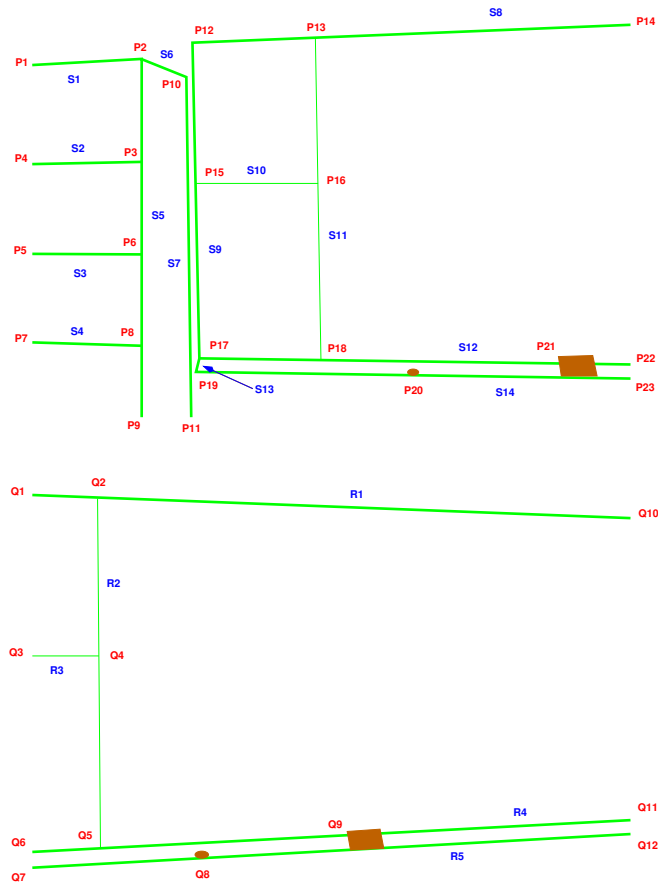


Figura 4:

En concret, si designem per P_i els punts del primer esquema i per Q_j els del segon, d'acord amb la figura 4 on també es dona nom als segments que generen la imatge, i calculem les seves coordenades sobreposant-los a una quadrícula d'una determinada escala² s'obtenen les taules de coordenades següents:

²En el programa de dibuix que s'ha utilitzat per a calcar les línies rectes, les dues fotografies mesuraven 15 cm. d'amplada i s'han assignat les coordenades començant a comptar en centímetres des del punt més baix i més a l'esquerra possible seguint dues direccions perpendiculars.

Punt	P1	P2	P3	P4	P5	P6
Coord. X	0,0000	2,7422	2,7422	0,0000	0,0000	2,7422
Coord. Y	8,8311	8,9845	6,4022	6,3467	4,0956	4,0822
Punt	P7	P8	P9	P10	P11	P12
Coord. X	0,0000	2,7422	2,7422	3,8622	3,9844	4,0156
Coord. Y	1,8756	1,7911	0,0000	8,5289	0,0000	9,3911
Punt	P13	P14	P15	P16	P17	P18
Coord. X	7,0978	15,0000	4,0911	7,1667	4,1889	7,2378
Coord. Y	9,5089	9,8400	5,8622	5,8622	1,4756	1,4222
Punt	P19	P20	P21	P22	P23	
Coord. X	4,1044	9,5489	13,1933	15,0000	15,0000	
Coord. Y	1,1356	1,1267	1,5200	1,3222	0,9689	
Punt	Q1	Q2	Q3	Q4	Q5	Q6
Coord. X	0,0000	1,6356	0,0000	1,6689	1,7089	0,0000
Coord. Y	9,3467	9,2578	5,3111	5,3111	0,4800	0,3956
Punt	Q7	Q8	Q9	Q10	Q11	Q12
Coord. X	0,0000	4,2489	7,8956	15,0000	15,0000	15,0000
Coord. Y	0,0000	0,3267	0,9667	8,7667	1,1934	0,8445

Un cop es tenen aquestes dades s'intenta determinar una transformació que converteixi els punts (les coordenades) d'una de les fotografies en els punts corresponents de l'altra. Concretament, determinem els coeficients d'una matriu

$$M = \begin{pmatrix} M_{11} & M_{12} & M_{13} \\ M_{21} & M_{22} & M_{23} \\ M_{31} & M_{32} & M_{33} \end{pmatrix}$$

tals que la transformació φ donada per

$$\varphi(x, y) = \left(\frac{M_{11}x + M_{12}y + M_{13}}{M_{31}x + M_{32}y + M_{33}}, \frac{M_{21}x + M_{22}y + M_{23}}{M_{31}x + M_{32}y + M_{33}} \right)$$

envii quatre dels punts Q_j sobre els quatre punts P_i corresponents.

Si es consideren les parelles de punts corresponents

$$\begin{aligned} P_{13} &\leftrightarrow Q_2 \\ P_{16} &\leftrightarrow Q_4 \\ P_{20} &\leftrightarrow Q_8 \\ P_{21} &\leftrightarrow Q_9 \end{aligned}$$

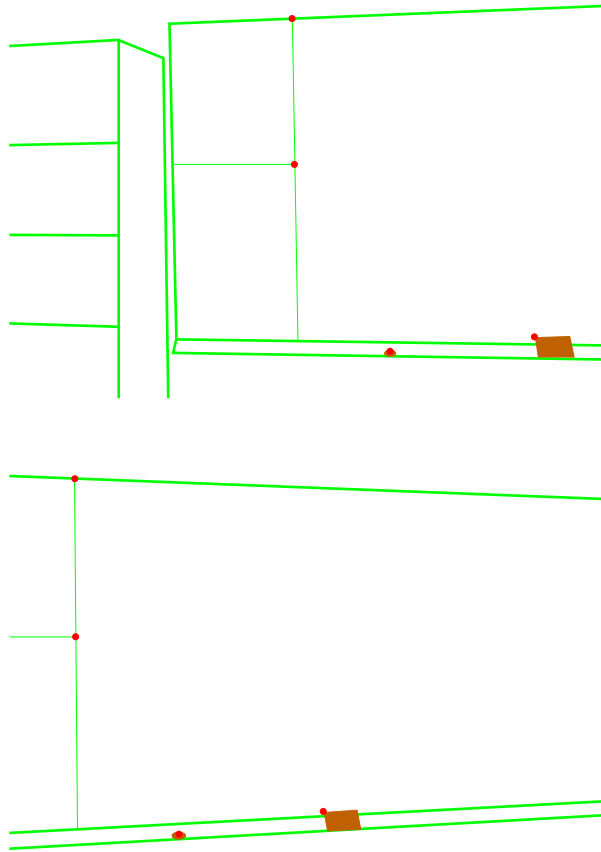


Figura 5:

representats per punts vermells en la figura 5 les equacions que s'han de resoldre són

$$7.0978 = \frac{1.6356 M_{11} + 9.2578 M_{12} + M_{13}}{1.6356 M_{31} + 9.2578 M_{32} + M_{33}}$$

$$9.5089 = \frac{1.6356 M_{21} + 9.2578 M_{22} + M_{23}}{1.6356 M_{31} + 9.2578 M_{32} + M_{33}}$$

$$7.1667 = \frac{1.6689 M_{11} + 5.3111 M_{12} + M_{13}}{1.6689 M_{31} + 5.3111 M_{32} + M_{33}}$$

$$5.8622 = \frac{1.6689 M_{21} + 5.3111 M_{22} + M_{23}}{1.6689 M_{31} + 5.3111 M_{32} + M_{33}}$$

$$9.5489 = \frac{4.2489 M_{11} + 0.3267 M_{12} + M_{13}}{4.2489 M_{31} + 0.3267 M_{32} + M_{33}}$$

$$1.1267 = \frac{4.2489 M_{21} + 0.3267 M_{22} + M_{23}}{4.2489 M_{31} + 0.3267 M_{32} + M_{33}}$$

$$9.5489 = \frac{4.2489 M_{11} + 0.3267 M_{12} + M_{13}}{4.2489 M_{31} + 0.3267 M_{32} + M_{33}}$$

$$1.1267 = \frac{4.2489 M_{21} + 0.3267 M_{22} + M_{23}}{4.2489 M_{31} + 0.3267 M_{32} + M_{33}}$$

Que, com es pot veure al primer cop d'ull, donen un sistema d'equacions lineals homogeni respecte de les incògnites M_{ij} .

Amb una mica d'ajuda d'algun programa de càlcul simbòlic (o fins i tot a ma amb una mica de paciència) es pot obtenir com a solució del sistema

$$\begin{aligned} M_{11} &= -885.5416953 M_{32} & M_{12} &= 18.89559819 M_{32} & M_{13} &= -6978.200410 M_{32} \\ M_{21} &= 94.06168005 M_{32} & M_{22} &= -1068.435875 M_{32} & M_{23} &= -1317.205184 M_{32} \\ M_{31} &= 18.10934107 M_{32} & M_{32} &= \text{lliure} & M_{33} &= -1201.443408 M_{32} \end{aligned}$$

on es pot decidir que el paràmetre lliure M_{32} val 1 i obtenir valors concrets per a tots els paràmetres.

Un cop assignats aquests valors als paràmetres M_{ij} es pot determinar la transformació que produeix φ a tots els punts Q_j i als segments de l'esquema corresponent³. Dibuint la figura que apareix després de fer aquesta transformació (vermell) juntament amb la figura que determinen els punts P_i (verd) s'obté la figura 6, on es pot comprovar l'alt grau de coincidència entre les parts comuns als dos esquemes. En particular, la prolongació dels extrems superior i inferior de la pissarra queda força alineada.

Si en comptes d'haver buscat una transformació com φ s'hagués intentat fer el mateix amb una *transformació afí* del tipus

$$\psi(x, y) = (ax + by + u, cx + dy + v)$$

³És important remarcar que aquest tipus de transformacions converteix segments en altres segments i que, per tant, tota la informació necessària per a reconstruir la imatge d'un d'aquests esquemes són les coordenades dels punts extrems de les línies que hi apareixen.

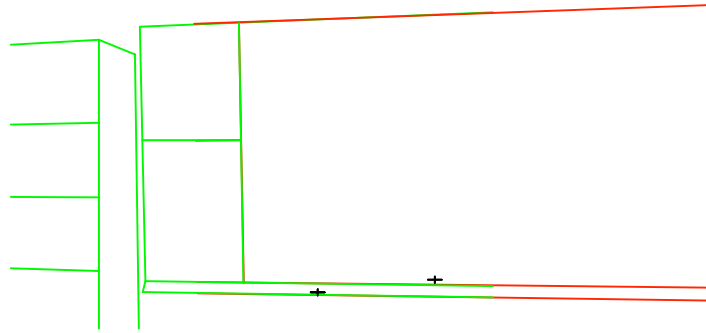


Figura 6:

(d'aquest tipus són, en particular, tots els moviments rígids del pla, és a dir girs, translacions, simetries i les seves composicions, i totes les homotècies) només es podrien haver posat tres condicions en comptes de quatre i el resultat que s'obtidria considerant la correspondència

$$\begin{aligned} P_{13} &\rightsquigarrow Q_2 \\ P_{16} &\rightsquigarrow Q_4 \\ P_{20} &\rightsquigarrow Q_8 \end{aligned}$$

seria el de la figura 7 que és clarament molt més dolent que l'anterior ja que

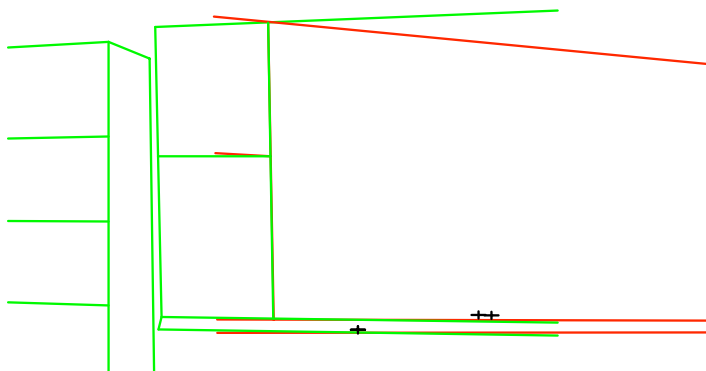
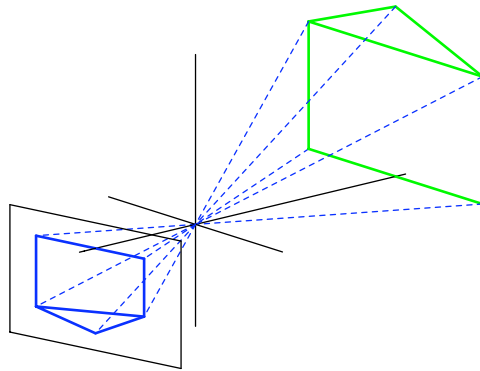


Figura 7:

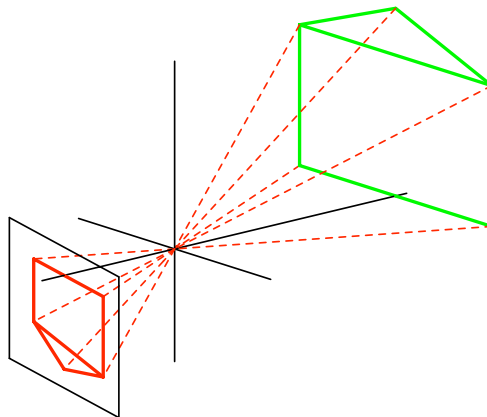
alineja la part inferior de la pissarra però no ho fa amb la part superior.

I com és que funciona?

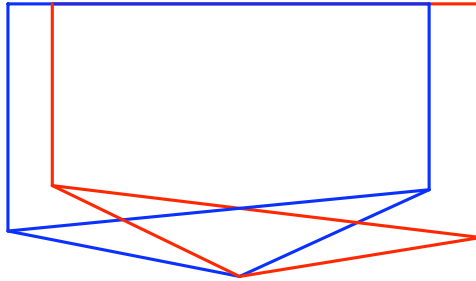
Quan es fa una fotografia els raigs de llum que provenen del tema passen tots pel focus de l'objectiu i acaben fent una marca en la pel·lícula fotogràfica o en el sensor digital. Això vol dir que quan es mira el que hi ha en el negatiu fotogràfic o el que detecta el sensor de la màquina digital es veu una representació dels raigs que passen per aquest punt que determina el focus de l'objectiu. Esquemàticament (i sempre que l'objecte sigui pla o estigui prou lluny), es pot pensar que el procés d'obtenir una fotografia és com el que es representa en la figura següent



on la figura verda és l'objecte que es fotografia i la blava la imatge que apareix projectada en la càmera. Quan es canvia el punt de vista de la càmera la projecció resultant és diferent i apareix una imatge com la vermella de l'esquema següent

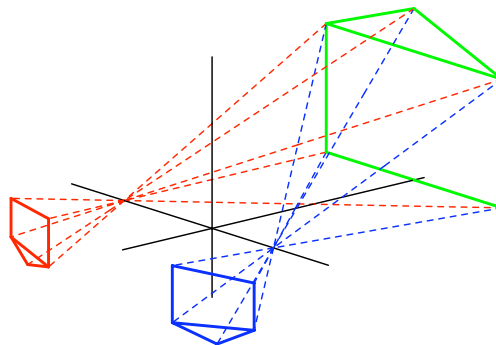


Representant les dues figures una al costat de l'altra es pot comprovar que, tot i ser representacions del mateix objecte, són ben diferents entre si.



Què és, doncs, el que tenen en comú les dues imatges obtingudes? Una resposta possible a aquesta pregunta és que cada cop que es pren una fotografia es fa una representació, cada cop sobre un pla diferent, del conjunt de rectes de l'espai \mathbb{R}^3 que passen pel focus⁴ i, per tant, que aquest procés és una manera *d'assignar coordenades* a aquest conjunt de rectes. Així que les dues figures només són les representacions del mateix conjunt de rectes *en dos sistemes de coordenades diferents*. Des d'aquest punt de vista el problema s'ha convertit en descobrir quin és el *canvi de coordenades* que es produeix quan es fa un canvi en el pla de projecció.

Un cop d'ull a l'esquema següent hauria de deixar clar que fins i tot en el cas en que el focus de la projecció canvia de lloc el problema continua sent del mateix tipus i que els càlculs que s'hauran de fer seran els mateixos.



⁴En realitat hi ha tot un conjunt de rectes de \mathbb{R}^3 , les que són paral·leles al pla sobre el que es fa la projecció, que no es poden representar sobre un pla. Aquestes rectes no representen un problema per a la situació que s'estudia ja que la seva aparició seria equivalent a treballar amb un objectiu amb un angle de visió proper a 180 graus.

Les coordenades homogènies

Des del punt de vista de \mathbb{R}^3 , les rectes que passen per l'origen queden determinades en el moment que es dóna el seu *vector director* encara que en aquesta assignació hi ha un grau d'indeterminació provinent del fet que si s'està considerant la recta amb vector director $\vec{v} = (a, b, c)$ també serà vector director de la mateixa recta qualsevol múltiple de \vec{v} de la forma $\lambda\vec{v} = (\lambda a, \lambda b, \lambda c)$, mentre s'agafi $\lambda \neq 0$. Això fa que quan parlem d'assignar coordenades a les rectes que passen per l'origen s'ha de precisar que és el mateix dir que es pren la recta (a, b, c) que dir que es pren la recta $(\lambda a, \lambda b, \lambda c)$ per a qualsevol $\lambda \neq 0$. Quan es fa això es diu que s'han donat *les coordenades homogènies* d'una d'aquestes rectes⁵.

Les coordenades homogènies depenen de l'elecció d'una base de \mathbb{R}^3 amb la indeterminació produïda pel fet que les bases $\vec{e}_0, \vec{e}_1, \vec{e}_2$ i $\lambda\vec{e}_0, \lambda\vec{e}_1, \lambda\vec{e}_2$, on λ és un coeficient qualsevol diferent de 0, donen les mateixes coordenades homogènies a cada una de les rectes de \mathbb{R}^3 que passen per l'origen. Quan es fixa una base de \mathbb{R}^3 i es calculen les coordenades homogènies en funció d'aquesta base (o de qualsevol base obtinguda multiplicant els mateixos tres vectors per un mateix coeficient) es diu que s'ha donat un *sistema de coordenades homogènies*.

És clar en aquest context que quan es fa un *canvi de coordenades homogènies* i una recta que tenia coordenades homogènies $X = (a, b, c)$ passa a tenir coordenades homogènies $X' = (a', b', c')$ hi ha una certa matriu 3×3 invertible M (la mateixa per a totes les rectes) per a la que es compleix $\lambda X' = (M \cdot X)$, on λ és un coeficient que depèn dels representats X i X' que s'hagin triat per a les coordenades homogènies de la recta en cada un dels sistemes.

Ara no és massa difícil comprovar que per a determinar la matriu d'un canvi de coordenades homogènies (que és única llevat de multiplicar tots els coeficients pel mateix factor) és necessari, i suficient, conèixer les coordenades de quatre rectes (que siguin *independents*) en els dos sistemes de referència. Sense entrar en detalls l'explicació és la següent: quan es planteja el sistema d'equacions que relaciona les coordenades homogènies de les quatre rectes en els dos sistemes de coordenades ($\lambda_i X'_i = (M \cdot X_i)$ per a $i = 1, 2, 3, 4$), el que

⁵Per al públic interessat en la Geometria Projectiva i, en particular, per les coordenades homogènies es pot recomanar el llibre de A. Reventós titulat *Geometria Projectiva* de la col·lecció "Materials" (no. 85) de la Universitat Autònoma de Barcelona o, també, el llibre (en anglès) *Projective Geometry* de A. Beutelspacher i U. Rosenbaum (Cambridge U.P.)

apareix és un sistema de 12 equacions lineals homogènies, tres equacions per cada recta (una per cada component), amb 13 incògnites, 9 dels coeficients d'una matriu 3×3 i un paràmetre de proporcionalitat λ_i per a cada recta, fet que en condicions d'independència deixa un grau de llibertat per a les solucions i determina la unicitat del canvi de coordenades homogènies. En canvi, si es pretén fixar el canvi de coordenades amb un nombre inferior de restriccions és clar que les solucions del sistema que apareix tenen més d'un grau de llibertat i per tant el canvi de coordenades homogènies no és únic. Per un altre costat, si s'intenten imposar més condicions apareixeran sistemes d'equacions homogenis amb més equacions que incògnites que, normalment, només acceptaran solucions en les que la matriu M no és invertible (en general l'única solució serà que M sigui la matriu nul·la) i, per tant, no dona un canvi de coordenades.

Per a tancar el cercle la pregunta que cal posar-se és: quina relació hi ha entre les coordenades homogènies i les coordenades que s'obtenen prenent com a representació de cada recta el punt on talla un pla determinat?

Per a respondre només s'ha de veure que, quan es fixa un sistema de referència dins d'un pla i es calculen coordenades respecte d'aquesta referència, hi ha un únic sistema de coordenades homogènies respecte del que les coordenades homogènies d'una recta que passa per l'origen i talla al pla en un punt al que li assignem coordenades (α, β) són $(\alpha, \beta, 1)$ o, recíprocament, si es considera la recta de coordenades homogènies $(a, b, c) \sim (a/c, b/c, 1)$ les coordenades *sobre el pla* del punt d'intersecció són $(a/c, b/c)$.

A partir d'aquí es veu que quan es prenen mesures sobre dues imatges diferents de la mateixa escena hi haurà una matriu 3×3 invertible que transformarà les coordenades homogènies corresponents a una representació en les corresponents de l'altra. Per a conèixer aquesta matriu serà suficient calcular les coordenades, respecte dels dos sistemes de referència, de quatre punts que surtin en les dues imatges ja que els sistemes de coordenades homogenis queden determinats quan es fixen les coordenades de quatre rectes.

De forma explícita, si es tenen quatre punts P_1, P_2, P_3 i P_4 , visibles en dues imatges de la mateixa escena, tals que en les coordenades d'una de les imatges són

$$P_i = (\alpha_i, \beta_i)$$

i en les coordenades de l'altra

$$P_i = (\alpha'_i, \beta'_i)$$

hi ha un canvi de coordenades homogènies, determinat per una matriu M de la forma

$$M = \begin{pmatrix} M_{11} & M_{12} & M_{13} \\ M_{21} & M_{22} & M_{23} \\ M_{31} & M_{32} & M_{33} \end{pmatrix}$$

tal que

$$\begin{pmatrix} M_{11} & M_{12} & M_{13} \\ M_{21} & M_{22} & M_{23} \\ M_{31} & M_{32} & M_{33} \end{pmatrix} \cdot \begin{pmatrix} \alpha_i \\ \beta_i \\ 1 \end{pmatrix} = \lambda_i \begin{pmatrix} \alpha'_i \\ \beta'_i \\ 1 \end{pmatrix}$$

(per a $i = 1, 2, 3, 4$). Tenint en compte que aquest sistema es pot reescriure com un grup de quatre sistemes de la forma

$$\begin{aligned} \alpha_i M_{11} + \beta_i M_{12} + M_{13} &= \alpha'_i \lambda_i \\ \alpha_i M_{21} + \beta_i M_{22} + M_{23} &= \beta'_i \lambda_i \\ \alpha_i M_{31} + \beta_i M_{32} + M_{33} &= \lambda_i \end{aligned}$$

i que el valor dels paràmetres λ_i és irrellevant per a la solució del problema, es pot aïllar cada una de les incògnites λ_i de l'última equació de cada grup i deixar el sistema com quatre grups d'equacions de la forma

$$\begin{aligned} \frac{\alpha_i M_{11} + \beta_i M_{12} + M_{13}}{\alpha_i M_{31} + \beta_i M_{32} + M_{33}} &= \alpha'_i \\ \frac{\alpha_i M_{21} + \beta_i M_{22} + M_{23}}{\alpha_i M_{31} + \beta_i M_{32} + M_{33}} &= \beta'_i \end{aligned}$$

que són, exactament, les equacions que s'han utilitzat per a resoldre el problema en el cas de les dues imatges de la pissarra.

Galeria d'imatges

A continuació podeu veure unes quantes imatges una mica més entretingudes que la pissarra obtingudes a partir de fotografies fetes amb una càmera digital compacta no massa sofisticada i retocades amb el programa d'ús lliure anomenat `hugin` que es pot descarregar des de <http://hugin.sourceforge.net/>. Aquest programa té com a punt de partida l'algoritme que hi ha explicat a les seccions anteriors. Naturalment, realitza moltes més operacions dedicades a minimitzar les distorsions produïdes pels errors de mesura en les posicions

dels punts triats; amagar el millor possible la zona on se sobreposen dues, o més, imatges; corregir les distorsions sistemàtiques del tipus de lent de la càmera fotogràfica; i moltes més opcions que podeu descobrir si feu un cop d'ull a la pàgina mencionada anteriorment.

En cada cas es mostren les fotografies originals i el resultat obtingut després del procés de retoc. És probable que en cada una d'elles detecteu d'una forma més o menys clara el lloc per on s'ha realitzat la *cosidura* entre les fotografies (detall que amb una mica de calma es pot anar refinant per a fer que sigui gairebé imperceptible), però el que és més remarcable és el fet que es pot modificar el punt de vista de les fotografies per aconseguir que es vegin totes dues sota el mateix angle de visió (que si us fixeu, pot ser que no coincideixi amb l'angle de visió de cap de les imatges parcials).





I, per acabar, un parell de paisatges d'Irlanda.





Departament de Matemàtiques
Universitat Autònoma de Barcelona
gguasp@mat.uab.cat

Publicat el 5 d'octubre de 2007