

# Repte 5

Alumne: Martí Caixal Joaniquet  
NIU: 1563587, data: 22/03/2022

## 1 Introducció

Al repte 5 es demana eliminar reflexes amb homografies. S'ha de fer un seguit de fotos a una superfície amb reflexes. L'activitat es divideix en dos apartats principals:

- Homografia amb les imatges
- Eliminació del reflexe

## 2 Imatges utilitzades

Per poder treballar correctament amb les imatges, cada una d'elles ha de tenir el reflexe a una posició diferent. En aquest cas, s'ha realitzat 4 fotografies canviant la posició i orientació tant de la càmera com de la font de llum.

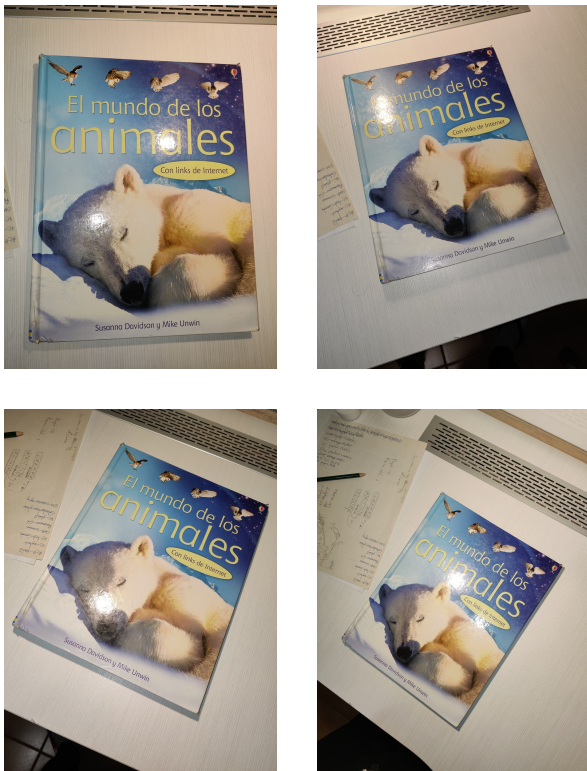


Figura 1: Imatges originals

A la figura 1 es veu com les zones amb reflexes són diferents. Més endavant es veurà per què és important que els reflexes no es sobreposin a les diferents imatges.

Si bé l'activitat es podria realitzar amb simplement 2 fotografies, més endavant es veurà la raó d'utilitzar-ne més.

## 3 Punts de control

Per poder realitzar els passos següents, cal trobar uns punts de referència comuns entre les imatges. Amb el

típus d'objecte que estem tractant i les transformacions que caldrà fer, els punts de control poden ser perfectament les 4 cantonades del llibre.

Cada un d'aquests punts està representat per un sistema de coordenades en dos dimensions. El codi que hi ha a continuació mostra com s'obtenen:

```
1 imshow(im1);  
2 x1 = []; y1 = [];  
3 for j = 1:4  
4     zoom on;  
5     pause();  
6     zoom off;  
7     [x1(j), y1(j)] = ginput(1);  
8     zoom out;  
9 end
```

De totes maneres, com que s'ha de fer moltes execucions i proves, els valors obtinguts es guarden a un arxiu i posetriorment es llegeixen mitjançant:

```
1 save('im1', 'x1', 'y2');  
2 load im1.mat
```

## 4 Transformacions

Un cop es té els punts de referència de cada imatge, es pot començar a buscar les transformacions necessàries que porten les diferents imatges a la primera d'elles. Al tractar-se d'una superfície plana, la transformació que porta un pla a un pla és l'anomenada homografia.

El mètode per obtenir-la entre parelles de 4 punts o més és la DLT (Direct Linear Transform). El codi utilitzat per crear la matriu i solucionar el sistema és el següent:

```
1 M = [];  
2 for i = 1:4  
3     M = [  
4         M;  
5         x1(i) y1(i) 1 0 0 0 -x2(i)*x1(i)  
6         -x2(i)*y1(i) -x2(i);  
7         0 0 0 x1(i) y1(i) 1 -y2(i)*x1(i)  
8         -y2(i)*y1(i) -y2(i)  
9     ];  
10 [u, s, v] = svd(M);  
11 H = reshape(v(:,end), 3, 3)';  
12 H = H / H(3,3);
```

Per portar la imatge cap a la primera, s'ha d'aplicar la homografia inversa. El gran inconvenient és que per defecte es preserva la fotografia completa (veure figura 2). Si bé el llibre té la transformació correcta respecte orientació i posició, no comparteix les mateixes mides amb l'altre figura



Figura 2: Comparació transformació simple

Per solucionar el problema, cal generar un marc de referència per la primera imatge i, posteriorment, passar-la com a paràmetre a la resta de transformacions:

```
1 RO = imref2d(size(im1));
2 proj = projective2d(inv(H'));
3 im2C = imwarp(im2, proj, 'OutputView',
  RO);
```

La figura 3 mostra les 3 imatges transformades correctament per quadrar amb la primera, ahora que també s'aconsegueix que el llibre tingui la mateixa mida a totes.

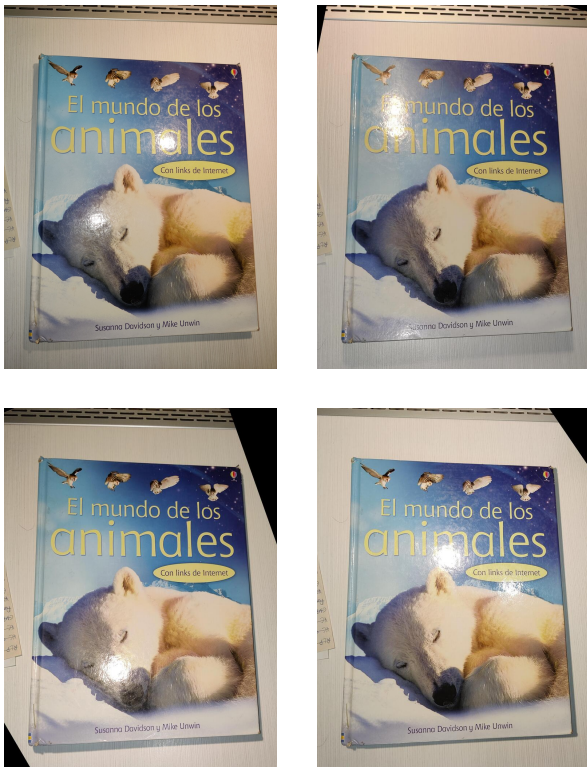


Figura 3: Imatges transformades correctament

## 5 Eliminació de reflexes

Ara que ja es té el llibre de totes les imatges amb la mateixa posició, orientació i mida, es pot procedir al pas final d'eliminar els reflexes. Primer, però, és convenient

guardar les 4 imatges en una sola estructura de dades per poder operar amb comoditat:

```
1 imArrCorrected = zeros(im_height,
  im_width, im_channels, 4);
2 imArrCorrected(:, :, :, 1) = im1;
3 imArrCorrected(:, :, :, 2) = im2C;
4 imArrCorrected(:, :, :, 3) = im3C;
5 imArrCorrected(:, :, :, 4) = im4C;
```

Les dos operacions que han donat el millor resultat són:

1. Mediana: consisteix en fer la mediana píxel a píxel entre les 4 imatges. És a dir, el resultat de cada píxel serà al valor que més es repeteixi entre les 4 imatges
2. Mitjana: consisteix en calcular la mitjana dels diferents valors d'un mateix píxel.
3. Mínim: consisteix en quedar-se amb el valor més petit que té un mateix píxel a les múltiples imatges.

Per calcular els resultats es fa ús del següent codi:

```
1 med = median(imArrCorrected, 4);
2 avg = mean(imArrCorrected, 4);
3 minim = min(imArrCorrected, [], 4);
```

### 5.1 Resultats utilitzant mediana

La figura 4 mostra els resultats obtinguts mitjançant una operació amb la mediana.

El primer resultat és utilitzant únicament 2 imatges. Pel propi comportament de la mediana, amb únicament 2 imatges no es pot aconseguir un resultat satisfactori i es veuen clarament els reflexes d'ambdues imatges.

El segon resultat utilitza 3 imatges i ara sí que es veu una gran millora. Tot i això, encara es pot apreciar algun que altre reflexe al mosell de l'animal, zona entre els ulls i el nom de l'autor. Com que ara es té 3 imatges per comparar, i els reflexes estan tots a zones diferents, mitjançant la mediana sempre guanyen els valors repetits, és a dir, els que no tenen reflexes.

Per últim, hi ha el resultat utilitzant les 4 imatges, sense rastre dels pocs reflexes que encara existien al resultat anterior. Tant utilitzant 3 com utilitzant 4 imatges, els reflexes estan sempre a zones diferents, per tant, el funcionament de la mediana implica que s'hagi d'obtenir el mateix resultat (1 píxel amb reflexe v.s. N sense reflexe). Segurament, com que les fotos estan capturades canviant la font de llum, les tonalitats dels colors canvien mínimament entre foto i foto. Així doncs, com més fotos es tingui, més fàcil serà trobar realment dos píxels que siguin idèntics a les 3 capes de colors. Aquesta és la raó que permet millors resultats utilitzant 4 imatges en comptes de 3.

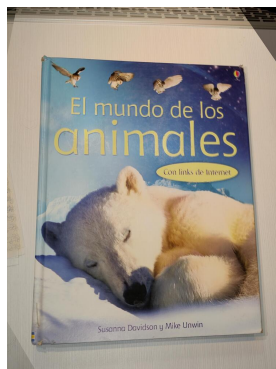
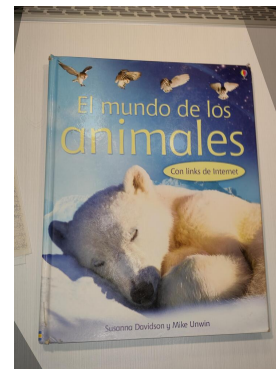
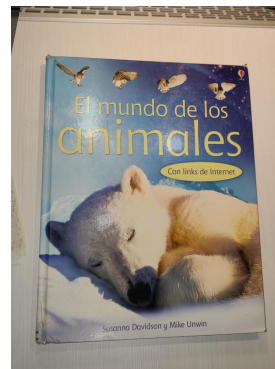
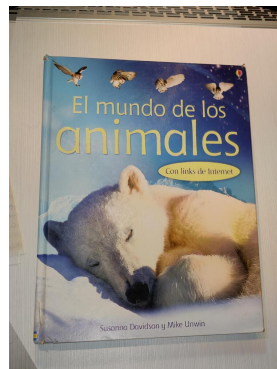
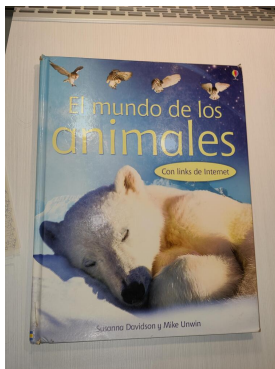


Figura 4: Mediana amb 2, 3 i 4 imatges

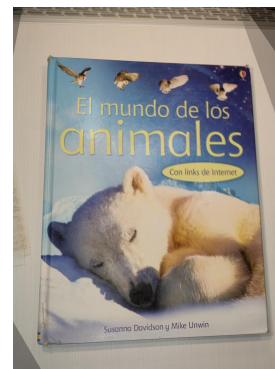


Figura 5: Mitjana amb 2, 3 i 4 imatges

## 5.2 Resultats utilitzant la mitjana

La figura 5 mostra els resultats obtinguts mintjançant una operació amb la mitjana. Es veu que, tot i millorar a mesura que s'utilitzen més imatges, els resultats queden lluny dels aconseguits amb la mediana. Al cap i a la fi, amb la mitjana sempre és manté un percentatge del reflex.

Si per exemple s'utilitzen 3 imatges, el valor de cada píxel serà un 33% del de cada imatge. Així, la zona amb reflexes manté una tercera part del reflexe. Per tal d'aconseguir un bon resultat, caldria tenir bastantes més imatges per així minimitzar la participació que tenen els reflexes al resultat final.

## 5.3 Resultats utilitzant el mínim

La figura 6 mostra els resultats obtinguts mintjançant una operació amb el mínim.

Si bé no s'ha obtingut els millors resultats, sí s'ha obtingut uns resultats molt interessants si es comparen amb els altres dos mètodes. Els reflexos es veuen principalment blancs, és a dir, els valors que defineixen els colors de la zona són molt grans. Per tant, fent una operació de mínim ens asegurem que mai s'agafa un reflexe, a no ser que el mateix píxel per totes les fotos tingui un reflexe. Utilitzant dues fotos, sempre que els reflexos estiguin en zones diferents, hi hauria d'haver suficient per eliminar-los. El problema, però, és que les fotos tenen diferents il·luminacions i sempre es queda amb el píxel més fosc. Es pot observar que, a mesura que s'utilitzen més imatges, la fotografia sencera es va enfosquir.

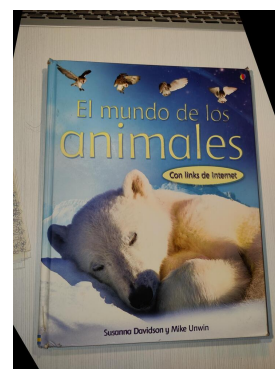
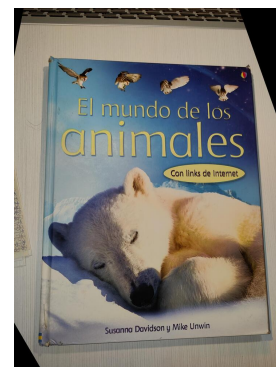
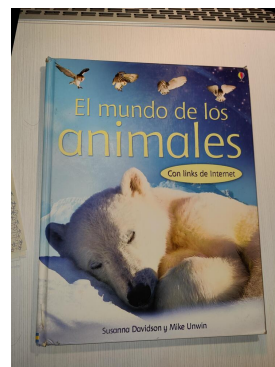


Figura 6: Mínim amb 2, 3 i 4 imatges



## 6 Resultat final

El millor resultat s'ha obtingut utilitzant una operació de mediana amb 4 imatges. La figura 7 mostra la imatge resultant final.

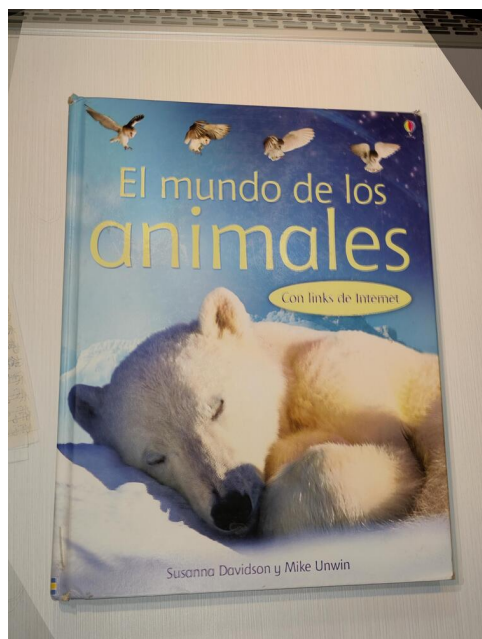


Figura 7: Millor resultat