

Réalisation d'un panorama

LUDOVIC LAMARCHE
QUENTIN PERALES
ELIE POUSSOU

E.I.S.T.I
PAU

19 janvier 2014

Table des matières

| | | |
|----------|--|----------|
| 1 | Les six fonctions du deuxième livrable | 3 |
| 1.1 | Grayscale | 3 |
| 1.2 | Binnary | 4 |
| 1.3 | Histogramme | 4 |
| 1.4 | Dilate | 5 |
| 1.5 | Erode | 6 |
| 1.6 | Convolution [wikiConvo] | 7 |
| 2 | Notre algorithme d'automatisation d'un panorama | 9 |
| 2.1 | Récupération des points clé | 9 |
| 2.2 | Algorithme de comparaison des points clé | 14 |
| 2.3 | Fonction de collage des images | 14 |
| 2.4 | Notre algorithme principal | 14 |
| 2.5 | Comment améliorer cet algorithme? | 14 |

Introduction

Le projet se conclue par ce dernier livrable en fin de semestre. L'objectif est donc de créer un programme qui assemble des images ou des photos pour réaliser un panorama. Pour cela, il a fallu mettre en oeuvre toutes nos connaissances accumulées au cours des deux premiers mois, ainsi que nos recherches pour remplir l'objectif fixé par ce projet.

Dans un premier temps, nous allons présenter les fonctions réalisées pour le second librable car elles ont été très utiles pour la réalisation de notre algorithme codé durant cette dernière partie.

Dans un second temps, nous allons présenter les étapes ainsi que les fonctions principales de notre algorithme qui relie les images entre elles. Puis, nous expliquerons les améliorations que nous aurions pu apporter au programme.

Chapitre 1

Les six fonctions du deuxième livrable

1.1 Grayscale

La fonction Grayscale permet de transformer une image couleur en teintes de gris.

Dans une image en couleur, chaque pixel est codé par 3 composantes (Rouge, Vert et Bleu).

En grayscale, chaque pixel sera codé par une seule composante.

La formule utilisée est la suivante :

$$pixelGrayscale = 0,299 * pixelRouge + 0,587 * pixelvert + 0,114 * pixelbleu$$

On parcourt donc l'image couleur, et on applique cette formule à chaque pixel, c'est à dire à ces trois composantes.

Cette fonction effectue donc une moyenne pondérée des pixels rouges, verts et bleus.

L'image obtenue sera donc 3 fois plus petite, et en noir et blanc.

1.2 Binnary

Chaque pixel d'une image binaire ne peut avoir pour valeurs que 0 ou 1. La fonction binary nécessite un seuil et une image en teintes de gris, le seuil étant choisi par l'utilisateur.

On parcourt l'image en teintes de gris, et pour chaque pixel, si son intensité est supérieure au seuil, on lui alloue la valeur 0 qui correspond à la couleur blanche. Si l'intensité du pixel est supérieure au seuil, on lui alloue la valeur 1 qui correspond à la couleur noire.

Voici l'algorithme utilisé :

```
1   POUR TOUT les pixels de l'image
    SI le pixel >seuil
        pixel=1
    FINSI
5   FINPOUR
```

1.3 Histogramme

L'histogramme d'une image permet d'illustrer la répartition des teintes d'une image.

Le but est de compter pour chaque intensité le nombre de pixels qui sont de cette intensité.

D'abord on crée un tableau de taille égale à la teinte maximale de l'image dont on veut obtenir l'histogramme.

Ensuite, on parcourt l'image et pour chaque pixel, on rajoute 1 dans la case du tableau qui correspond à l'intensité de ce pixel.

On peut résumer tout ceci grâce à l'algorithme suivant :

```
1 tab : tableau d'entiers de taille égale à la teinte maximale de  
1 'image  
  
3 POUR chaque pixel de l'image  
tab [teinte_pixel]++  
5 FINPOUR
```

On peut choisir de tracer cet histogramme mais le tableau suffit.

1.4 Dilate

Pour la dilatation nous avons utilisé un algorithme qui utilise des "objects pixel" tiré du livre Machine Vision[[livre](#)]. Un objet pixel est un groupe de neuf pixels formant un carré de taille 3x3 et qui respecte certaines conditions. Pour simplifier l'algorithme on nomme chaque case de 0 à 8.

| | | |
|---|---|---|
| 3 | 2 | 1 |
| 4 | 8 | 0 |
| 5 | 6 | 7 |

Si on utilise les conventions de logique combinatoire, c'est à dire que le signe . signifie ET et le signe + signifie OU, les conditions d'un objet pixel se résume comme ceci :

$$\begin{aligned} & 8.[((1+2+3).5+6+7).\bar{4}.\bar{0}] \\ & +[(1+0+7).(3+4+5).\bar{2}.\bar{6}] \\ & +[3.(5+6+7+0+1).\bar{2}.\bar{4}] \\ & +[1.(3+4+5+6+7).\bar{2}.\bar{0}] \\ & +[7.(1+2+3+4+5).\bar{0}.\bar{6}] \\ & +[5.(7+\circ+1+2+3).\bar{4}.\bar{6}] \end{aligned}$$

Finallement, il suffit d'appliquer l'algorithme suivant :

```
1 POUR TOUT les pixels de l'image
    SI le pixel est un objet pixel ALORS
        copier le groupe de pixels dans l'image de destination
    FINSI
5 FINPOUR
```

Étant donné que cette méthode enlève beaucoup de pixels, il n'est pas nécessaire de faire une érosion en suivant. Cependant nous avons utilisé un autre algorithme moins restrictif pour l'érosion.

1.5 Erode

Pour l'érosion nous utilisons un masque. Nous testons un groupe de pixel et nous assignons la valeur 1 si tout les pixels dans le masque sont à 1. Sinon on retourne un 0. Nous avons pris directement l'algorithme d'opération pixel[operationpixel] :

```
1 destination : matrice de l'image érodée
    source : matrice de l'image à éroder
3 x,y,k1,k2 : variables d'incrémentation

5 POUR CHAQUE pixel de l'image FAIRE //x et y étant la position
    POUR k1 allant de -1 à 1 FAIRE
        POUR k2 allant de -1 à 1 FAIRE
            destination[x][y] = destination[x][y] ET source[x+k1][y+k2]
9     FINPOUR
    FINPOUR
11 FIN POUR
```

1.6 Convolution [wikiConvo]

La convolution permet d'appliquer un filtre de convolution à une image en échelle de gris. Le but est d'affecter pour chaque pixel de l'image des coefficients aux pixels autour de celui-ci.

La taille du filtre est variable, cependant, elle est toujours paire. Les filtres les plus utilisés sont les filtres de taille 3x3, mais parfois, les filtres de tailles supérieures sont requis. Les filtres sont récupérés en début de fonction, une erreur est générée si le filtre n'est pas correct, c'est à dire si le fichier texte comporte des caractères autres que des chiffres ou nombres ou s'il n'y a pas assez de coefficients.

Afin d'expliquer l'algorithme qui mène à la convolution, nous allons utiliser un filtre 3x3 sur une image. Prenons le filtre F suivant :

0 1 0

1 1 1

0 1 0

F est appliqué à un pixel p de coordonnées (x, y). La valeur du pixel après la convolution sera alors

$$\begin{aligned} nouvelleValeur = & 0 * p(x - 1, y - 1) + 1 * p(x, y - 1) + 0 * p(x + 1, y - 1) \\ & + 1 * p(x - 1, y) + 1 * p(x, y) + 1 * p(x + 1, y) \\ & + 0 * p(x - 1, y + 1) + 1 * p(x, y + 1) + 0 * p(x + 1, y + 1) \end{aligned}$$

De plus, on crée un décalage pour éviter d'assigner des coefficients à des pixels hors image. Pour un filtre 3x3, le décalage est de 1 ; pour 5x5, il est de 3 ; pour 7x7, il est de 5. On peut donc dire que

$$decalage = \frac{taille - 1}{2} \quad (1.1)$$

Voici un exemple de convolution que l'on peut obtenir :

$$\begin{array}{ccccccccc} 35 & 40 & 41 & 45 & 50 & & 35 & 40 & 41 & 45 & 50 \\ 40 & 40 & 42 & 46 & 52 & 0 & 1 & 0 & 35 & 40 & 41 & 45 & 50 \\ 42 & 46 & 50 & 55 & 55 & X & 0 & 0 & 0 & 40 & 40 & 42 & 46 & 52 \\ 48 & 52 & 56 & 58 & 60 & 0 & 0 & 0 & 42 & 46 & 50 & 55 & 55 \\ 56 & 60 & 65 & 70 & 75 & & & & 56 & 60 & 65 & 70 & 75 \end{array}$$

Après avoir vu certaines fonctions très utiles dans le traitement d'image, nous allons maintenant voir comment les enchaîner pour les rendre utiles lors de la création du panorama.

Chapitre 2

Notre algorithme d'automatisation d'un panorama

Dans cette partie, nous allons expliquer plusieurs fonctions principales dans notre algorithme.

Tout d'abord, la récupération des points clés. Ensuite, la comparaison de ces points clés entre les images. Puis, la fonction de collage qui permet d'assembler deux images connaissant le décalage. Enfin, expliquer l'enchaînement de ces fonctions pour aboutir à notre programme.

Pour finir cette partie, nous expliquerons les améliorations que nous pourrions apporter à notre code.

2.1 Récupération des points clé

La récupération des points clés est une étapes cruciale dans la réalisation d'un panorama car ils permettent de savoir ce que l'image a de particulier, et de connaître les formes qu'on pourrait comparer avec une autre image.

Pour cela, nous avons mis en oeuvre plusieurs algorithmes que nous avons trouvé lors de nos recherches. Trois algorithmes majeurs ont été trouvés.

Tout d'abord, la transformée de Hough devait nous permettre de récupérer les points issus de rencontre de droites ou de cercles. Cependant, lors de la comparaison de formes compliquées, les points trouvés n'étaient pas assez

précis et suffisants pour permettre une comparaison efficace.

Puis, nous avons codé le décodeur de Harris qui s'appuie sur des calculs mathématiques, notamment les matrices pour récupérer des points clés. Ce détecteur n'a pas toujours données des points pouvant être utilisés.

Ensuite, nous avons commencé à coder la méthode SIFT, une méthode très efficace, mais très longue.

Entre temps, nous avons trouvé un autre moyen pour trouver des points clé, qui utilisait les fonctions expliquées dans la partie précédente.

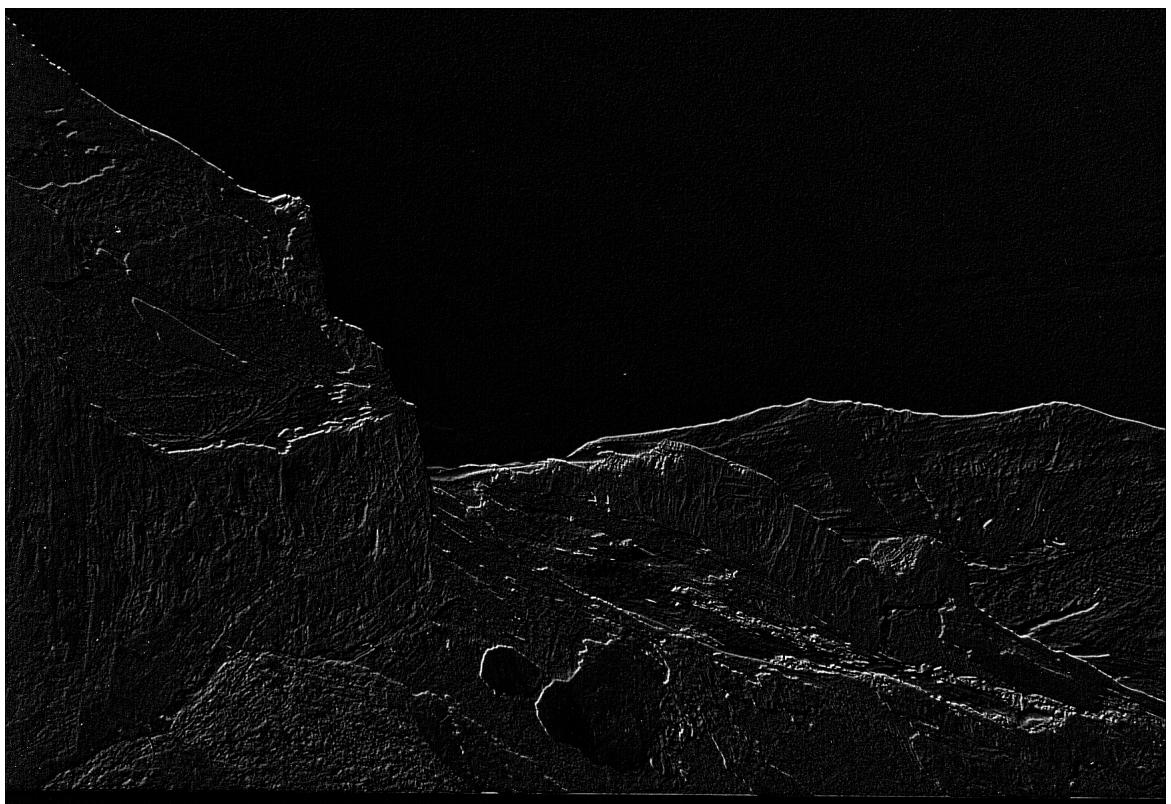
Raisonnons à partir de cette image :



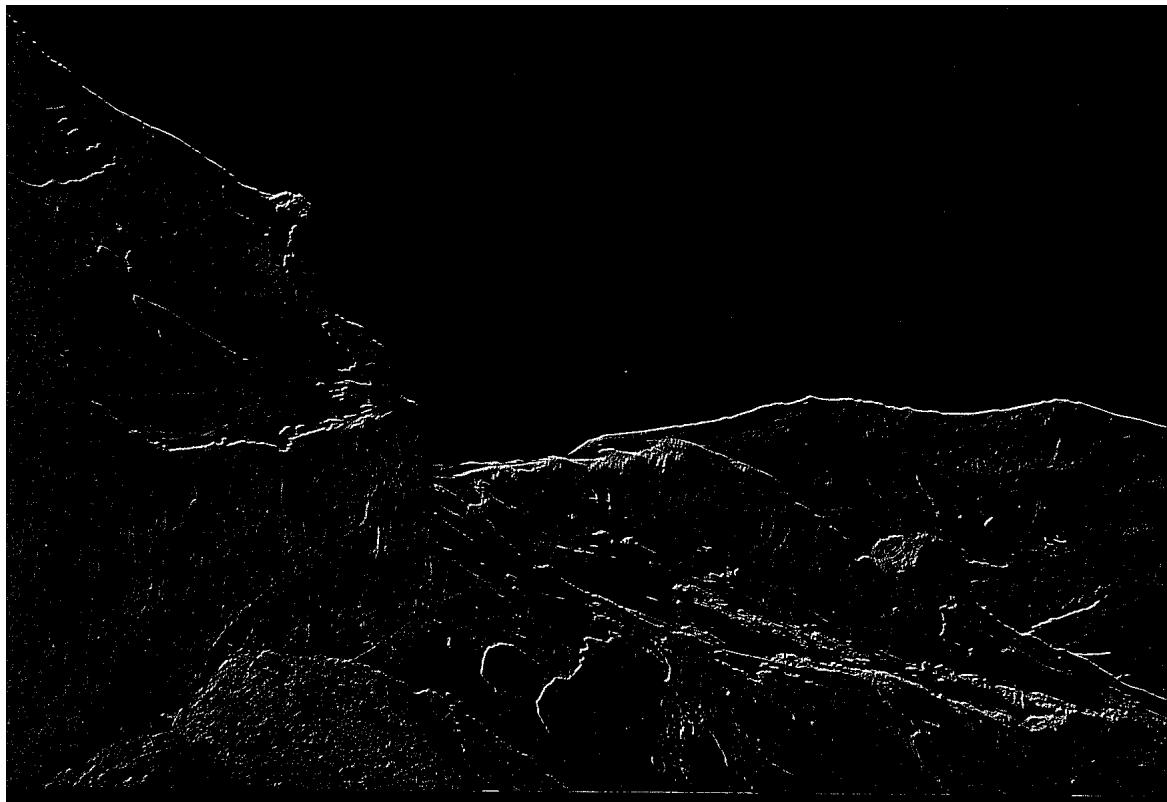
Pour cela, à partir d'une image en couleur, il faut la passer en greyscale puis effectuer une convolution avec un filtre de Prewitt.

Cela permet de détecter les contours tout en limitant l'effet du bruit sur la photo.

$$\begin{array}{ccc} 1 & 1 & 1 \\ \text{filtre de Prewitt : } & 0 & 0 & 0 \\ & -1 & -1 & -1 \end{array}$$



Ensuite, on converti cette image en binaire avec un seuil de 100.



Enfin, on opère à une dilatation puis une érosion cinq fois de suite pour enlever les points esseulés.



Pour finir cette détection, on parcourt l'image pour voir si un point blanc possède quinze voisins dans un carré de 15x15 pixels autour de lui. Cela peut s'apparenter à une forme de dilatation. Cette dernière opération est très utile lors de l'algorithme de comparaison des points clés expliqué ci-dessous.



2.2 Algorithme de comparaison des points clé

2.3 Fonction de collage des images

2.4 Notre algorithme principal

2.5 Comment améliorer cet algorithme ?

Conclusion