

PONTIFICIA UNIVERSIDAD CATÓLICA DEL PERÚ
FACULTAD DE CIENCIAS E INGENIERÍA



PONTIFICIA
UNIVERSIDAD
CATÓLICA
DEL PERÚ

**ADAPTACIÓN DE LOS ALGORITMOS SIFT Y LSH PARA LA
DIFERENCIACIÓN DE ARCHIVOS DE IMÁGENES**

Tesis para optar el Título de **Ingeniero Informático**, que presenta las bachilleres:

Tania Gabriela Ramirez Franco
Ila Ibañez Quispe

ASESOR: Mg. Claudia Zapata del Rio

Lima, Mayo de 2016



Anexo 1: Detalles de pruebas

Para poder obtener la curva Precision/Recall se tomó una muestra de 52 imágenes del dominio tomado de [1]. Se tomó como muestra máxima dicha cantidad debido a que la información de parecido no era suficiente para poder realizar una comparación con los resultados del algoritmo (no había información de las relaciones entre todas las imágenes). En la figura 1 se muestra un collage de las 52 imágenes tomadas.



Figura 1: Collage de imágenes utilizadas. Adaptado de [1]

El sitio del dominio además, ofrece información del parecido de las imágenes. Esta
Tania Gabriela Ramírez Franco — Ila Ibañez Quispe Página III

información se distribuye en documentos que siguen la siguiente clasificación:

- 1) Good(bueno).- Indica que la imagen es una foto clara del objeto
- 2) Ok(bien).- Indica que más del 25 % del objeto es visible
- 3) Bad(malo).- Indica que el objeto no está presente
- 4) Junk(basura).- Indica que menos del 25 % del objeto es visible

A pesar de haber 4 clases, la documentación solo tenía información de la distribución de 3 de ellas (good, ok y junk) por lo cual se decidió trabajar con estas 3. La información estaba distribuida de la siguiente forma: Existía un total de 55 imágenes “Query”, en base a las cuales se realizaba la clasificación de las otras imágenes. Por cada imagen “Query” existían 4 archivos de texto, el primero correspondía al archivo que contenía el nombre del archivo “Query” mientras que los otros 3 eran los archivos que contenían las listas de imágenes clasificadas como “good”, “ok” o “junk”.

En base a dichas listas y con el nombre del archivo “Query” se procedió a armar un cuadro de doble entrada, como el que devuelve el algoritmo, que muestra la relación entre todas las imágenes. El cuadro quedó como se muestra en la figura 2.

Como se puede apreciar en la figura 2 no toda la tabla pudo ser llenada con la información del sitio. Por tanto, se decidió que en los casos en los que no hubiera información de la relación entre 2 imágenes, se tomaría como valor el asignado a la categoría Junk. Tomando esto en cuenta, se terminó con el llenado de la tabla.

Una vez lista, se pudo comparar esta tabla base contra las tablas generadas por el algoritmo. Esto se logró utilizando las diferentes herramientas del programa Microsoft Excel, las cuales permitieron realizar de forma rápida y eficiente la comparación entre tablas. Finalmente, se obtuvieron los datos para poder obtener los valores necesarios para la curva Precision/Recall.

Anexo 2: Pruebas de validación previas

A continuación se describen las pruebas que se realizaron a la herramienta pero que fueron descartadas debido a que se consideraron muy subjetivas.

Metodología

Para poder sustentar la precisión de la solución elaborada, se realizó un estudio para contrastarla contra la visión humana. Este estudio consistió en la elaboración de encuestas (ver detalle en el anexo 1) cuyos resultados han sido comparados con los resultados obtenidos con la herramienta.

En las encuestas se presentaron 10 pares de imágenes y se les preguntó a los encuestados si es que estas eran iguales, parecidas o diferentes.

El algoritmo, por su parte, permite obtener un resultado de “En que porcentaje se parecen las imágenes”.

Con estos resultados, la experimentación pretende demostrar que el algoritmo es más preciso que la visión humana.

Variables

Las variables que se utilizaron en las pruebas fueron las siguientes:

Media de aciertos por individuo: Se refiere al promedio de aciertos por cada individuo que realizó la encuesta.

Media de aciertos por par de imágenes: Se refiere al promedio de aciertos por cada par de imágenes comparadas.

Pruebas de Hipótesis

A continuación se muestran las pruebas de hipótesis formuladas para la experimentación.

Prueba de Hipótesis 1: Precisión de similitud por individuo

Lo que se busca demostrar con esta prueba es la diferencia de la precisión entre el algoritmo y una persona. Por tanto, se pretende demostrar que la media de aciertos del algoritmo es mayor a la media de aciertos que tuvieron los encuestados.

Sean:

MAA: Media de aciertos del algoritmo.

MAH: Media de aciertos de una persona.

Ho: $MAA = MAH$: La media de los aciertos del algoritmo es igual a la media de aciertos obtenidos por una persona.

Ha: $MAA > MAH$: La media de los aciertos del algoritmo es mayor a la media de aciertos obtenidos por una persona.

Prueba de Hipótesis 2: Precisión de similitud por comparación

En este experimento, lo que se busca demostrar es la diferencia de la precisión entre el algoritmo y una persona por cada par de imágenes comparadas. Por tanto, se pretende demostrar que la media de aciertos del algoritmo es mayor a la media de aciertos que tuvieron los encuestados.

Sean:

MAIA: Media de aciertos del algoritmo.

MAIH: Media de aciertos de una persona.

Ho: MAIA = MAIH: La media de los aciertos del algoritmo es igual a la media de aciertos obtenidos por una persona.

Ha: MAIA > MAIH: La media de los aciertos del algoritmo es mayor a la media de aciertos obtenidos por una persona.

Estadístico a utilizar

Como el tamaño de la muestra utilizada es menor a 30, se recomienda utilizar el estadístico T-Student, de manera que se verifique la diferencia de medias. Para este experimento, se eligió un nivel de significación de 5 %, es decir hay una posibilidad de 0.05 de rechazar la hipótesis, siendo esta cierta.

Consideraciones para la experimentación

A continuación se especifican las consideraciones que se tuvieron por cada una de las pruebas de hipótesis.

Consideraciones para prueba de Hipótesis 1

Se realizó la encuesta a una población de 24 personas, por tanto, se utilizó una corrida del algoritmo con una proporción de 24 para poder hacer una comparación precisa.

Los datos obtenidos de las encuestas y la corrida del algoritmo se pueden apreciar en el anexo 1.

A continuación se muestra la tabla con los resultados procesados de ambos métodos.

Num	Aciertos del Encuestado	Aciertos del Algoritmo
1	7	3
2	6	6
3	8	5
4	6	7
5	9	7
6	8	6
7	8	7
8	7	5
9	5	6
10	8	6
11	7	5
12	5	4
13	7	5
14	8	7
15	5	8
16	8	6
17	6	6
18	7	6
19	5	6
20	8	8
21	5	7
22	6	6
23	4	4
24	6	6

Cuadro 1: Resultados de aciertos por cada individuo

En donde:

Aciertos del Encuestado: Total de aciertos por el encuestado.

Aciertos del Algoritmo: Total de aciertos por la corrida del algoritmo.

Consideraciones para prueba de Hipótesis 2

Se realizó la encuesta con un cantidad de 20 imágenes, distribuidas en 10 pares.

Los datos obtenidos de las encuestas y la corrida del algoritmo se pueden apreciar en el anexo 1.

A continuación se muestra la tabla con los resultados procesados de ambos métodos.

Num	Encuestado	Algoritmo
1	17	6
2	14	21
3	4	7
4	13	18
5	10	19
6	12	8
7	20	12
8	4	20
9	24	24
10	24	24

Cuadro 2: Resultados de aciertos por par de imágenes

En donde:

Encuestado: Cantidad de veces que el encuestado acertó en su respuesta.

Algoritmo: Cantidad de veces que el algoritmo acertó en su evaluación.

Resultados

A continuación se muestran los resultados obtenidos por cada prueba de hipótesis.

Resultados de la prueba de Hipótesis 1

Los supuestos para la aplicación de T-Student son el Supuesto de Normalidad y el Supuesto de Igualdad de Varianzas

Se utilizó la prueba de Shapiro-Wilk para verificar la hipótesis de que los datos provienen de una distribución normal

Hipótesis A: Supuesto de Normalidad

Ho: Los datos provienen de una distribución normal.

Ha: Los datos no provienen de una distribución normal.

En la tabla 3 se puede ver que los valores de significancia son 0,061 y 0,077. Como ambos son mayores a 0,05, no hay evidencia para rechazar ésta hipótesis nula. Por ende podemos afirmar con una seguridad del 95 % de que los datos provienen de una distribución normal, y por lo tanto podemos aplicar la diferencia de varianzas.

Anexo 2: Pruebas de validación previas

		Kolmogorov-Smirnov			Shapiro-Wilk		
		Estadístico	gl	Sig.	Estadístico	gl	Sig.
VAR	algoritmo	,180	24	,043	,921	24	,061
	encuestados	,236	24	,001	,925	24	,077

Cuadro 3: Pruebas de normalidad

Hipótesis B: Supuesto de Igualdad de Varianzas

Ho: No hay diferencia significativa entre las varianzas.

Ha: Hay diferencia significativa entre las varianzas.

		Prueba de Levene de calidad de varianzas	
		F	Sig.
VAR	Se asumen varianzas iguales	1,813	0,185
	No se asumen varianzas iguales		

Cuadro 4: Prueba de Levene para varianzas

Como el valor de significancia en este caso es de 0,185, tampoco hay evidencia para rechazar esta hipótesis nula, por lo que podemos aplicar la prueba de T-Student para obtener resultados más precisos.

Aplicando la prueba de T-Student obtenemos los siguientes resultados.

		Prueba T para la igualdad de medias						
		t	gl	Sig (bilateral)	Diferencia de medias	Diferencia de error estándar	95 % de intervalo de confianza de la diferencia	
							Inferior	Superior
VAR	Se asumen varianzas iguales	1,916	46	0,062	0,70833	0,36973	-0,03589	1,45256
	No se asumen varianzas iguales	1,916	45,515	0,062	0,70833	0,36973	-0,3611	1,45277

Cuadro 5: Prueba T-Student

Según la tabla 5 consideramos los resultados de la 1era fila, donde el valor de significancia resulta 0,031 (0,062/2), el cual es menor a 0,05. Por tanto, se concluye que hay suficiente evidencia para rechazar la hipótesis nula (de la hipótesis 1), quedándonos con la hipótesis alternativa:

Ha: $MAA > MAH$: La media de los aciertos del algoritmo es mayor a la media de aciertos obtenidos por una persona.

Resultados de la prueba de Hipótesis 2

Al igual que en el caso anterior, los supuestos para la aplicación de T-Student son el Supuesto de Normalidad y el Supuesto de Igualdad de Varianzas

Se utilizó la prueba de Shapiro-Wilk para verificar la hipótesis de que los datos provienen de una distribución normal

Hipótesis A: Supuesto de Normalidad

Ho: Los datos provienen de una distribución normal.

Ha: Los datos no provienen de una distribución normal.

		Kolmogorov-Smirnov			Shapiro-Wilk		
		Estadístico	gl	Sig.	Estadístico	gl	Sig.
VAR	algoritmo	,218	10	,197	,879	10	,126
	encuestados	,122	10	,200	,934	10	,486

Cuadro 6: Pruebas de normalidad

En la tabla 6 se puede ver que los valores de significancia son 0,126 y 0,486. Como ambos son mayores a 0,05, no hay evidencia para rechazar la hipótesis nula. Por ende podemos afirmar con una seguridad del 95% de que los datos provienen de una distribución normal y podemos aplicar la diferencia de varianzas.

Hipótesis B: Supuesto de Igualdad de Varianzas

Ho: No hay diferencia significativa entre las varianzas.

Ha: Hay diferencia significativa entre las varianzas.

		Prueba de Levene de calidad de varianzas	
		F	Sig.
VAR	Se asumen varianzas iguales	0,096	0,760
	No se asumen varianzas iguales		

Cuadro 7: Prueba de Levene para varianzas

Como el valor de significancia en este caso es de 0,760, tampoco hay evidencia para rechazar ésta hipótesis nula y podemos proceder a realizar la prueba de

T-Student para obtener un mejor resultado.

		Prueba T para la igualdad de medias						
		t	gl	Sig (bilateral)	Diferencia de medias	Diferencia de error estándar	95 % de intervalo de confianza de la diferencia	
							Inferior	Superior
VAR	Se asumen varianzas iguales	0,535	18	0,599	1,70000	3,17718	-4,97500	8,37500
	No se asumen varianzas iguales	0,535	17,989	0,599	1,70000	3,17718	-4,97530	8,37530

Cuadro 8: Prueba T-Student

Según la tabla 8 consideramos los resultados de la 1era fila, donde el valor de significancia resulta 0,2995 ($0,599/2$), el cual no es menor a 0,05. Por tanto, no hay suficiente evidencia para rechazar la hipótesis nula.

Sin embargo, podemos deducir que esto se debe a la comparación de los ciertos pares de imágenes, puesto que presentan factores (brillo, desenfoque, etc) que no permiten una evaluación muy precisa por el algoritmo SIFT, que es el que nos permite obtener los puntos de interés. Aun así, consideramos estos pares puesto que en un ambiente real es muy probable que se den estos casos.

Validez de la evaluación

En la experimentación se utilizó una cantidad de valores menor a 30, por lo cual se pudo utilizar el estadístico T-Student. Sin embargo, se estima que con una muestra mayor se podrían obtener resultados más precisos.

Pruebas de aceptación

A manera de evaluación propia, se realizó una prueba de comparación con la imagen de la figura 3.

Se tomó esta imagen pues posee diversos bordes y cambios de tonalidad los cuales son propicios para hallar puntos de interés. Luego, se buscaron imágenes similares para compararlas, recopilando las que se muestran en la figura 4.

Finalmente, se procedió a utilizar la herramienta para realizar las comparaciones y se obtuvieron los resultados que se muestran en la tabla 9.

En porcentajes los resultados se muestran como en la tabla 10.



Figura 3: Imagen base seleccionada

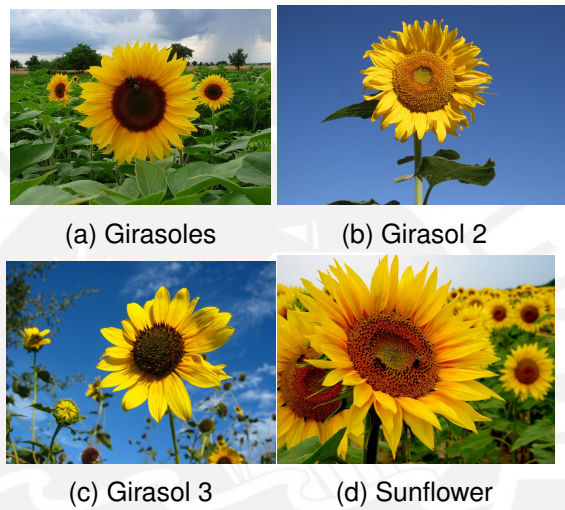


Figura 4: Imágenes similares a la imagen base

	girasol	girasoles	girasol-2	girasol-3	sunflower
girasol	-	0,681496	0,661017	0,667620	0,735738
girasoles	0,681496	-	0,664360	0,649465	0,671830
girasol-2	0,661017	0,664360	-	0,651216	0,684416
girasol-3	0,667620	0,649465	0,651216	-	0,661748
sunflower	0,735738	0,671830	0,684416	0,661748	-

Cuadro 9: Resultados prueba aceptación (Valores)

En base a nuestra análisis y experiencia, llegamos a la conclusión de que los resultados son acertados y que la herramienta esta funcionando correctamente.

	girasol	girasoles	girasol-2	girasol-3	sunflower
girasol	-	68,15 %	66,10 %	66,76 %	73,57 %
girasoles	68,15 %	-	66,44 %	64,95 %	67,18 %
girasol-2	66,10 %	66,44 %	-	65,12 %	68,44 %
girasol-3	66,76 %	64,95 %	65,12 %	-	66,18 %
sunflower	73,57 %	67,18 %	68,44 %	66,18 %	-

Cuadro 10: Resultados prueba aceptación (Porcentajes)

Anexo 3: Resumen de realización de encuestas

Se desarrollaron 2 encuestas para poder comparar la precisión de la visión humana contra la solución implementada (algoritmo). En un inicio se pensó elaborar una sola encuesta, pero al ver que los resultados de esta ya no eran comparables con los obtenidos con el algoritmo, se decidió realizar una segunda encuesta.

Primera encuesta

Elaboración

Se elaboró la encuesta enfocándola en obtener el tiempo que le tomaba a una persona el encontrar las diferencias entre 2 imágenes. Se consideraron 3 pares de imágenes, de manera que se percibieran los niveles de dificultad: fácil, moderado y difícil.

Muestra

Para esta encuesta se contó con 9 personas, pues al tratarse de una encuesta presencial no se pudo realizar con una mayor cantidad de personas. Todos los encuestados fueron estudiantes y egresados de ingeniería informática con edades entre los 20 y 30 años.

Resultados

Se obtuvo:

Anexo 3: Resumen de realización de encuestas

- Tiempo que le tomó a cada persona evaluar las imágenes
- Cantidad de diferencias encontradas

Lo cual se puede apreciar en la siguiente tabla

	Imagen 1		Imagen 2		Imagen 3	
	Tiempo	Coincidencias	Tiempo	Coincidencias	Tiempo	Coincidencias
Encuestado 1	1:00	7	1:00	6	2:00	10
Encuestado 2	1:00	7	1:00	6	2:00	10
Encuestado 3	1:00	7	2:00	6	2:00	9
Encuestado 4	0:50	7	1:00	7	2:00	14
Encuestado 5	1:00	6	1:00	6	2:00	8
Encuestado 6	1:00	7	2:00	6	2:00	7
Encuestado 7	1:00	6	1:00	6	2:00	12
Encuestado 8	1:00	7	1:00	7	3:00	12
Encuestado 9	1:00	7	3:00	7	2:00	15
Promedio	0:59	7	1:27	6	2:07	11
Diferencias existentes		7		7		15

Cuadro 11: Resultados de encuesta 1

Nota: Lo que se denominó como “Diferencias existentes” se refiere a la cantidad de diferencias que existen entre las imágenes utilizadas.

Resultados de la solución (algoritmo)

Los resultados obtenidos con el algoritmo son porcentajes que indican qué tan similares son ambas imágenes.

Conclusión en base a los resultados

Luego de ver los resultados de las encuestas y los del algoritmo, se observó que estos no eran comparables, puesto que la primera busca diferencias entre las imágenes y el segundo busca las similitudes. Por tanto, se decidió realizar una segunda encuesta.

Segunda Encuesta

Elaboración de la Encuesta

Se hizo una lista de 10 páginas, mostrando en cada una de ellas 2 imágenes. En cada página se realizaron las siguientes preguntas:

Tania Gabriela Ramírez Franco — Ila Ibañez Quispe

Página xvii

Pregunta: Observa las imágenes por X segundos. Son iguales o no?

Opciones de respuestas

- Si
- No

Pregunta: Son similares?

Opciones de respuestas

- Si
- No

Previamente se hicieron las siguientes indicaciones a los encuestados acerca de las definiciones utilizadas: Iguales: Que ambas imágenes son idénticas Diferentes: Que hay diferencias en las imágenes Similares: Que, a pesar de ser diferentes, las imágenes se parecen

Además, en cada página se especificaba el tiempo que el encuestado debía tomar para observar ambas imágenes, tiempo que es equivalente al que le toma al algoritmo en evaluar ambas imágenes. Esto se hacía con el fin de tener una base temporal constante.

Muestra

Para esta encuesta se contó con 24 personas puesto que esta se realizó en línea. En esta ocasión participaron todos los encuestados de la encuesta anterior más un grupo de estudiantes y egresados de otras especialidades(educación, psicología, comunicación, entre otros) cuyas edades están entre los 20 y 30 años.

Resultados de la Encuesta

Se obtuvieron las percepciones de los encuestados con respecto a las imágenes presentadas:

Anexo 3: Resumen de realización de encuestas

Pag 1	Pag 2	Pag 3	Pag 4	Pag 5	Pag 6	Pag 7	Pag 8	Pag 9	Pag 10
1	0	2	2	0	2	1	2	2	2
2	1	2	2	1	0	0	0	2	1
1	0	2	1	0	1	0	0	2	2
1	1	1	1	0	1	0	0	2	2
1	1	2	1	0	0	0	0	2	2
2	0	1	1	0	0	0	0	2	2
1	1	2	1	0	0	0	0	2	2
2	1	2	2	2	0	0	0	2	2
1	1	2	0	0	0	0	0	2	2
2	1	2	1	1	1	0	0	2	2
1	2	2	1	1	1	1	0	2	2
2	1	2	2	0	1	0	0	1	2
2	1	2	2	1	1	1	1	2	2
1	0	2	1	1	0	0	0	2	2
1	1	1	1	1	1	1	1	2	2
1	1	2	2	1	1	0	0	2	2
1	0	2	1	0	0	0	0	2	2
1	0	2	0	1	0	0	0	2	2
1	0	2	1	0	0	0	0	2	2
1	1	1	2	0	0	0	1	2	2
1	1	2	1	1	1	0	0	2	2
2	1	2	2	1	1	0	1	2	2
1	0	2	0	0	1	0	2	2	2
1	0	2	1	0	0	0	0	2	2

Cuadro 12: Resultados de encuesta 2

Cuya leyenda es:

2 = Las imágenes son iguales

1 = Las imágenes son similares

0 = Las imágenes son diferentes

Resultados de la solución

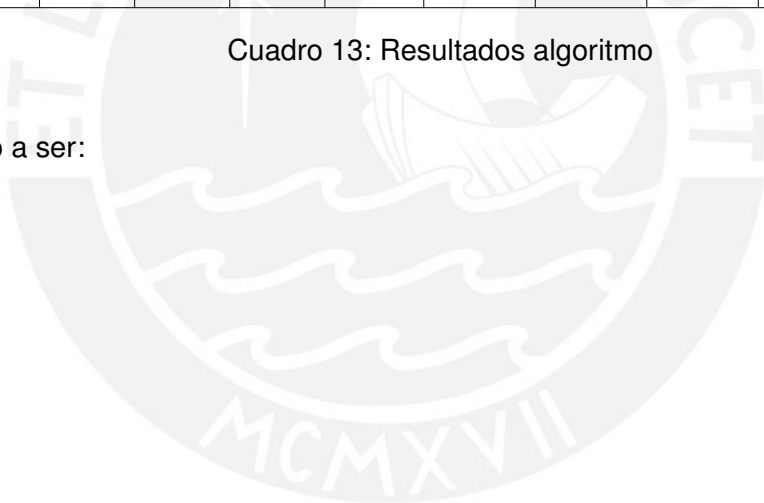
Los resultados del algoritmo, al igual que en el caso de la primera encuesta, son porcentajes que indican qué tan similares son ambas imágenes. Se utilizó el algoritmo para evaluar las imágenes utilizadas en la segunda encuesta obteniéndose los porcentajes correspondientes. Dichos porcentajes se ajustaron para poder tener resultados equivalentes a los obtenidos en la segunda encuesta. Lo que en un inicio era:

Anexo 3: Resumen de realización de encuestas

	1 vs 2	3 vs 4	5 vs 6	7 vs 8	9 vs 10	11 vs 12	13 vs 14	15 vs 16	17 vs 18	19 vs 20
1	0.1677	0.4022	0.2472	0.7183	0.3093	0.2141	0.1516	0.1566	1	1
2	0.2906	0.3486	0.2040	0.8855	0.5667	0.3932	0.3511	0.3491	1	1
3	0.2067	0.3701	0.2643	0.5878	0.1536	0.2380	0.2358	0.5528	1	1
4	0.2043	0.3906	0.1165	0.7263	0.4565	0.5998	0.5476	0.3550	1	1
5	0.2535	0.3691	0.3224	0.5320	0.5532	0.2769	0.2244	0.4815	1	1
6	0.2631	0.4980	0.2587	0.5327	0.2532	0.5374	0.4581	0.6155	1	1
7	0.1716	0.3815	0.0913	0.7879	0.2772	0.1776	0.1566	0.6627	1	1
8	0.2607	0.3665	0.2723	0.5473	0.3422	0.3323	0.3386	0.2038	1	1
9	0.1928	0.4614	0.1625	0.8466	0.6376	0.4521	0.4039	0.5072	1	1
10	0.2573	0.6312	0.1585	0.5709	0.2595	0.2849	0.2205	0.3475	1	1
11	0.2116	0.4120	0.1839	0.8719	0.7696	0.1712	0.1884	0.4214	1	1
12	0.2222	0.1835	0.1265	0.5850	0.7595	0.3543	0.2597	0.3340	1	1
13	0.2024	0.3665	0.1458	0.7828	0.5169	0.2715	0.2301	0.5609	1	1
14	0.2405	0.3532	0.1641	0.6862	0.2688	0.2231	0.1448	0.2938	1	1
15	0.2014	0.4234	0.2926	0.8517	0.2380	0.6083	0.5280	0.3314	1	1
16	0.3504	0.4413	0.2447	0.6532	0.3688	0.2750	0.2383	0.6279	1	1
17	0.1957	0.2245	0.1024	0.6243	0.1249	0.1422	0.0714	0.4445	1	1
18	0.1576	0.5162	0.2131	0.2938	0.2717	0.1776	0.1759	0.1614	1	1
19	0.1894	0.5737	0.1138	0.6425	0.1743	0.6567	0.5951	0.4177	1	1
20	0.2347	0.1220	0.3162	0.6331	0.4570	0.1926	0.1320	0.2681	1	1
21	0.2014	0.4143	0.1561	0.8984	0.3578	0.4795	0.4185	0.7458	1	1
22	0.1692	0.6688	0.3253	0.6070	0.2570	0.5689	0.5508	0.2032	1	1
23	0.1855	0.3553	0.1922	0.8573	0.1473	0.3832	0.3139	0.2799	1	1
24	0.1802	0.3274	0.2080	0.7285	0.3844	0.3997	0.3560	0.6402	1	1

Cuadro 13: Resultados algoritmo

Pasó a ser:



Anexo 3: Resumen de realización de encuestas

	1 vs 2	3 vs 4	5 vs 6	7 vs 8	9 vs 10	11 vs 12	13 vs 14	15 vs 16	17 vs 18	19 vs 20
1	0	1	0	1	1	0	0	0	2	2
2	1	1	0	2	1	1	1	1	2	2
3	0	1	1	1	0	0	0	1	2	2
4	0	1	0	1	1	1	1	1	2	2
5	1	1	1	1	1	1	0	1	2	2
6	1	1	1	1	1	1	1	1	2	2
7	0	1	0	1	1	0	0	1	2	2
8	1	1	1	1	1	1	1	0	2	2
9	0	1	0	2	1	1	1	1	2	2
10	1	1	0	1	1	1	0	1	2	2
11	0	1	0	2	1	0	0	1	2	2
12	0	0	0	1	1	1	1	1	2	2
13	0	1	0	1	1	1	0	1	2	2
14	0	1	0	1	1	0	0	1	2	2
15	0	1	1	2	0	1	1	1	2	2
16	1	1	0	1	1	1	0	1	2	2
17	0	0	0	1	0	0	0	1	2	2
18	0	1	0	1	1	0	0	0	2	2
19	0	1	0	1	0	1	1	1	2	2
20	0	0	1	1	1	0	0	1	2	2
21	0	1	0	2	1	1	1	1	2	2
22	0	1	1	1	1	1	1	0	2	2
23	0	1	0	2	0	1	1	1	2	2
24	0	1	0	1	1	1	1	1	2	2

Cuadro 14: Ajuste de resultados de algoritmo

Cuya leyenda es:

2 = El porcentaje es mayor a 80 %

1 = El porcentaje es mayor a 25 % y menor a 80 %

0 = El porcentaje es menor a 25 %

Cada fila de la tabla corresponde a una “tabla hash” generada por el algoritmo. Cada una de estas tablas evalúa las imágenes de forma diferente, de manera que a mayor cantidad de tablas habrá una mayor precisión en los resultados. La cantidad de estas tablas depende de la configuración establecida en el algoritmo. En este caso se estableció una cantidad de 15 para tener una mayor precisión sin que el tiempo de ejecución del algoritmo se vea afectado.

Anexo 4: Estructuras y formatos de imágenes

En esta sección se mostrarán las estructuras de los formatos de imagen a utilizar.

Estructura JPG

El formato JPG utiliza marcadores para denotar el inicio de un bloque determinado. Estos marcadores son de 2 bytes y empiezan de la forma '0xFF' seguidos del marcador del bloque siguiente.

Estructura GIF

GIF es diferente de muchos otros formatos de mapa de bits comunes en el sentido de que está basada en secuencias. Su estructura consiste en una serie de paquetes de datos, llamados bloques, junto con la información del protocolo adicional. Debido a esta disposición, los archivos GIF se deben leer como si fueran un flujo continuo de datos. Los diversos bloques y sub-bloques de datos definidos por el GIF se pueden encontrar casi en cualquier lugar dentro del archivo.

Estructura BMP

El archivo de mapa de bits se compone de estructuras de tamaño fijo (cabeceras), así como estructuras de tamaño variable que aparecen en una secuencia predeterminada. Contienen 4 secciones:

Anexo 4: Estructuras y formatos de imágenes

- Cabecera de archivo: Se usa para identificar al archivo, las aplicaciones suelen leer esta sección primero para asegurarse del tipo de archivo y de que no este dañado.
- Cabecera de imagen: Esta sección presenta información detallada de la imagen, como por ejemplo dimensiones, tamaño, resolución. La longitud de esta sección puede variar de acuerdo a la versión de bmp que este siendo utilizada, es por ello que contiene una variable indicando el tamaño de la cabecera.
- Tabla de colores: Es el listado de colores utilizados para representar la imagen. Estos son asociados a índices, los que figuran en la data de la imagen.
- Data de la imagen: Esta sección corresponde a los bits que representan la imagen

Estructura PNG

Para que la imagen tenga el formato png, sigue el siguiente proceso:

- Pase de extracción: Los píxeles de la imagen PNG se pueden reorganizar para formar varias imágenes más pequeñas llamadas imágenes o pases reducidos.
- Scanline serialización: la imagen se serializa en una línea de exploración a la vez. Los píxeles se ordenan de izquierda a derecha en una línea y las líneas se ordenan de arriba a abajo.
- Filtrado: cada línea de exploración se transforma en una línea de exploración por filtro usando uno de los metodos definidos para la compresión.
- Compresión: se produce en todas las líneas por filtro de la imagen filtrada.
- Fragmentación: la imagen comprimida se divide en partes de tamaño regular. Se añade un código de detección de errores para cada fragmento.
- Construcción de datastream: los trozos se insertan en el datastream.

Este formato se compone de:

Anexo 4: Estructuras y formatos de imágenes

- Cabecera de archivo: 8 bytes que indican que los siguientes fragmentos, corresponden a una sola imagen.
- Fragmentos: Cada uno de los fragmentos se compone de entre 3 a 4 campos: largo, tipo de fragmento, data del fragmento y CRC (data de validación)

Estructura PGM

PGM es el acrónimo para "Portable Gray Map", este formato representa una imagen en escala de grises. Es un formato no tan usado, pero sencillo y fácil de usar.

La estructura del archivo consta de:

- Un número mágico compuesto por 2 caracteres que identifican el tipo de archivo. Pueden ser P2 o P5
- Un espacio en blanco
- El ancho en formato decimal ASCII
- Otro espacio en blanco
- El alto en formato decimal ASCII
- Otro espacio en blanco
- El máximo valor para la escala de gris (Debe ser menor a 65536)
- Un cambio de línea
- Contenido de la imagen, que dependiendo del número mágico, pueden ser números entre 0 y el máximo valor para escala de gris especificado o expresado en formato binario compuesto por 1 o 2 bytes.

El algoritmo SIFT que se utiliza en el presente proyecto utiliza este tipo de archivo como entrada, por lo que todas las imágenes a procesar serán previamente convertidas a este formato.

Bibliografía

- [1] R. A. James Philbin and A. Zisserman. The oxford buildings dataset. [Online].
Available: <http://www.robots.ox.ac.uk/~vgg/data/oxbuildings/index.html>

