**INFORME CBIR PROCESSAT D’IMATGE**

En aquest informe presentem el nostre sistema per a la recuperació i identificació d’imatges similars a una imatge específica donada per l’usuari i que fa la posició de pregunta. El sistema utilitza descriptors definits per el standard MPEG7, en concret el model de color HMMD i el descriptor CSD (color structure descriptor). Per a validar els resultats utilitzem la base de dades U. Kentucky Benchmark Image Database, la qual conté 2000 imatges de 500 objectes diferents i té una gran varietat de colors i formes.

**1. INFORMACIÓ DEL DESCRIPTOR**

El procés d’extracció de característiques consisteix en dos parts. La primera consisteix a passar els fotos al espai de color HMMD i el segon pas és calcular el CSD (Color Structure Descriptor). A continuació expliquem en detall.

**Espai de color HMMD**

L’espai de color HMMD (Hue-Max-Min-Diff) és molt similar a un espai de color perceptualment uniforme. El cono doble limita el espai de color tal com és veu en la imatge posterior. Els components Max, Min, Diff estàn definits amb les següents equacions.

Es defineix també un altre parametre Sum.

I el Hue.

Llavors, tenim un total de cinc components en aquest espai de color, tot i així un conjunt de tres components {H, Max, Min} o {H, Diff, Sum} és suficient per a formar el espai de color HMMD i especificar un punt.

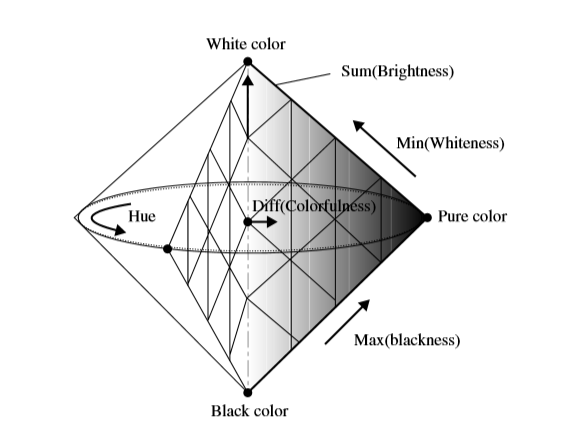


Figure 1- Representació de l'espai de color HMMD en doble cono

**Cuantificació del espai de color HMMD**

El espai de color HMMD és cuantifica de forma no uniforme per a utilitzarlo després amb el CSD.

És realitza una quantificació 3D de manera que es parteix l’espai en celes 3D. En l’estàndard MPEG7 és defineixes quatre quantificacions no uniformes. De 256, 128, 64, 32 celes respectivament.

Cada quantificació es defineix mitjançant 5 subespais HMMD de la següent manera. El aix del diff, té un rang [0,255] és talla en 5 sub-intervals: [0,6), [6, 20), [20, 60), [60, 110) and [110, 255). Aquesta partició 1D del eix – Diff implicitament defineix 5 subespais numerats del 0 al 4. Cada subespai és un subset del HMMD on el Sum i Hue poden prendre qualsevol valor.

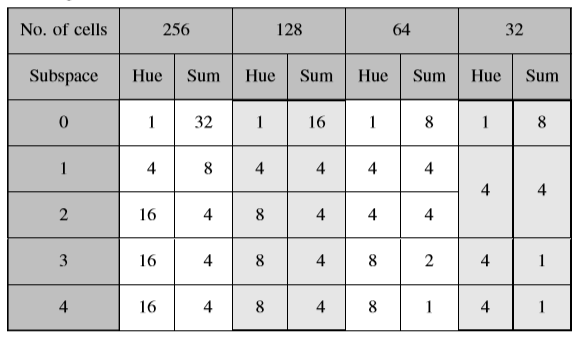


Figure 2- Quantificació del subespai HMMD per cada de les quatre particions

En la següent figura és veu un tall lateral en l’espai de color HMMD en el pla diff-sum pel angle de Hue 0 i mostra la quantificació en 128 celes. Els valors de l’eix Diff que determinen els límits del subespai están apuntats a la part superior. Les linees horitzontals de cada subespai ensenyen la divisió de l’eix sum en intervals uniformes.

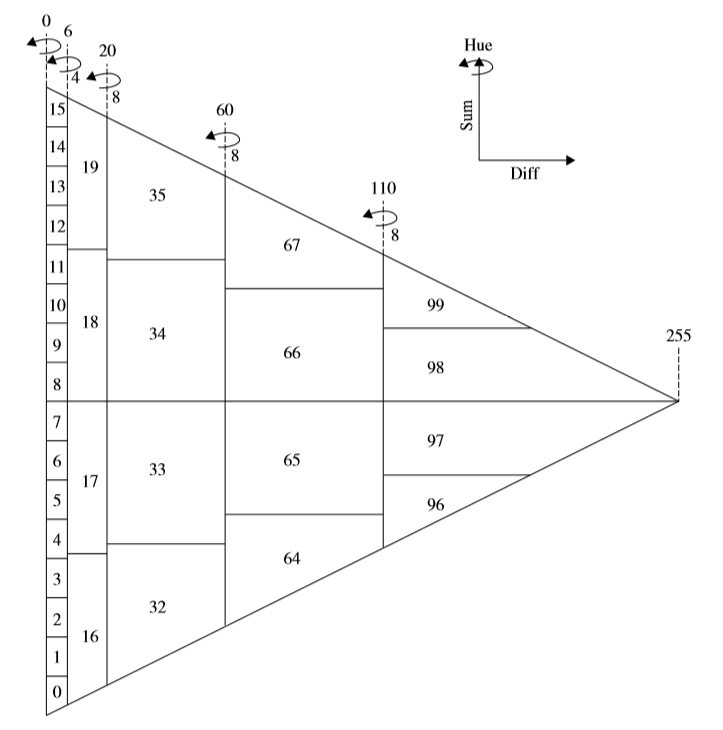


Figure 3- Tall del espai HMMD de 128 celes a hue = 0

**Color Structure Descriptor**

El color structure descriptor (CSD) representa una imatge mitjançant la distribució de color de l’imatge, similar al histograma de color i l’estructura espacial del color. L’informació adicional sobre l’estructura de color fa que el descriptor sigui sensible a certes caràcteristiques que l’histograma no detecta.

COMPLETAR EXPLICACIÓ

**2. ARQUITECTURA DEL SISTEMA**

El sistema ha sigut implementat seguint el següent esquema de blocs:

Lectura de fitxers de la carpeta:

S’utilitza la funció dir(src) per a llegir els fitxers d’una carpeta i ens retorna un struct amb els noms de totes les imatges de la base de dades. Això ens permetrà després anar recorrent cada imatge i calcular tot allò que sigui necessari.

Càlcul de la base de dades:

La base de dades està formada a partir dels histogrames de cada imatge i correspondrà en una matriu de 2000 files i tantes columnes com nivells s’hagin determinat.

Iterem sobre totes les fotos i realitzem els dos passos. El primer pas és convertir l’imatge al espai de color HMMD. Després calculem el CSD amb la funció hmmd\_histogram(photo, bins) on es passa com a arguments la matriu de la imatge que volem analitzar i el nòmbre de nivells de cuantificació que fem servir per a quantificar l’histograma.

El número de bins determinarà la quantitat de intervals en els que classificarem els píxels. Un número baix permet una gran rapidesa a l’hora de realitzar els càlculs però amb una baixa precisió. En canvi un número gran permet descriure una imatge amb fidelitat però amb un alt cost computacional.

Lectura del fitxer de input:

El sistema funciona utilitzant un arxiu d’entrada on es troben llistats els noms de les imatges de les quals volem posteriorment calcular els seus matches i les mètriques. És important determinar el número de fitxers per així després crear un vector d’histogrames de les imatges d’entrada.

Exemple de fitxer input:

ukbench01701.jpg

ukbench00926.jpg

ukbench01883.jpg

ukbench00116.jpg

ukbench00213.jpg

Lectura de la foto pregunta:

A partir d’aquest bloc fins a l’últim s’executaran per a cada foto llegida del fitxer de input. En aquest pas llegim la foto a la qual haurem de fer el match, i calculem el seu CSD.

Algoritme de matching:

L’algoritme de matching es basa en calcular la similitud entre l’histograma d'origen i aquells que es troben emmagatzemats en la base de dades. Per al càlcul d’aquestes distàncies s’han implementat diferents mètodes: chi2, bhattacharyya, intersection, kolmogorov, seuclidean,minkowski, mahalanobis.

Les distàncies han estat ordenades en sentit descendent i es retornen les 10 primeres posicions corresponents a les imatges més semblants.

Escriptura dels resultats:

Escriu els resultats retornats per l’algoritme de matching en un fitxer output indicant a quina imatge pregunta corresponen cadascuna. Exemple de fitxer output:

Retrieved list for query image ukbench01701.jpg

ukbench01701.jpg

ukbench01703.jpg

ukbench01700.jpg

ukbench01702.jpg

ukbench01808.jpg

ukbench00567.jpg

ukbench00638.jpg

ukbench00487.jpg

ukbench00639.jpg

ukbench00565.jpg

Retrieved list for query image ukbench00926.jpg

ukbench00926.jpg

ukbench00923.jpg

ukbench00922.jpg

...

**Càlcul de mètriques i gràfics:**

Calculem la precision i el recall per a cada nombre de resultats amb la següent formula:

On nums\_correctes es un array amb els noms de les imatges que corresponen al mateix objecte, aquests noms els calculem amb una Look Up Table (LUT) gràcies a que totes les imatges del mateix objecte estan agrupades en grups de 4. El procediment a seguir és el següent: primerament es fa el mòdul 4 del valor absolut de la foto, amb el

resultat s'obté la posició dintre del grup en la que es troba la nostra imatge de referencia, s’accedeix a la fila de la LUT corresponent i els valors de les columnes indiquen els valors a sumar per tal d’obtenir els valors i noms de les altres imatges que també són correctes.

Un cop es té el llistat d’imatges correctes i l’obtingut pel nostre algoritme s’ha de calcular la precision i el recall agafant el conjunt de totes les imatges correctes i en cada iteració un element més del conjunt calculat. D’aquesta manera s’obtenen 10 punts que conformen la funció de precison-recall.

La f-score ens indica en un sol valor la qualitat de les nostres mesures a partir dels valors de precision i recall. La calculem amb lasegüent comanda:

A continuació detallem aquestes mètriques i expliquem com s’avaluen els resultats.

**3.METODE D’AVALUACIÓ**

Per a avaluar els resultats s’han utilitzats tres paràmetres: la corva de precision-recall, la f-score i l’anàlisis qualitatiu. Utilitzant conjuntament aquestes mètriques hem analitzat i millorat aquest sistema fins arribar al punt actual.

**Corva de precision-recall**

Per a obtenir aquesta corva hem de calcular els valors de precision I recall per als resultats obtinguts i en funció del nombre de resultats. En la taula inferior hi ha un exemple de taule de precisió i recall.

TAULA PRECISION RECALL

Aquest és un mètode visual de determinar la qualitat de la nostra búsqueda, sent el resultat òptim una corba apegada a l’eix superior i dret. Per a crear el gràfic posem el recall en l’eix de les ordenades i la precision en l’eix de les abscisses. Depenent del tipus d’aplicació podria ser més interessant potenciar la precision, si volem que el nostre sistema seleccioni només imatges correctes, o el recall, si volem que es reconeguin totes les respostes correctes dintre del conjunt calculat tot i seleccionar-ne d’incorrectes.

FOTO FUNCIO PRECISION RECALL

**F-score**

És una mesura de precisió que té el nostres test. S’utilitza un valor ponderat de la precisió i del recall per a calcular la puntuació. Aquest mètode al retornar un únic valor és millor a l’hora de comparar moltes mètriques entre si. La fòrmula per a calcular-la és la següent.

**Anàlisis qualitatiu:**

Finalment per a analitzar els resultats obtinguts visualitzem les imatges trobades i determinem a ull si són correctes. La foto següent és un exemple de representació dels resultats:

Exemple de resultat visual mostrat per pantalla

FOTO EXEMPLES

**4. RESULTATS**

A continuació es mostren resultats que s’han aconseguit utilitzant els diferents mètodes de càlcul de similitud dels histogrames. La majoria consisteixen en utilitzar diferents mètriques ponderades per a calcular la distància

TAULES DE RESULTATS EN FUNCIO DE HISTOGRAMA NORMAL I CSD I EN FUNCIO DEL NOMBRE DE BINS

**5. TEMPS DE CÀLCUL**

TAULES DE TEMPS EN FUNCIO DE HISTOGRAMA NORMAL I CSD I EN FUNCIO DEL NOMBRE DE BINS (ES POT FER EL CÀLCUL DE 10 FOTOS I FER LA MITJANA)

EL MATCH DINS LA BASE DE DADES SEMPRE ES IGUAL

I TEMPS DE CÀLCUL FFT