

LINKÖPINGS UNIVERSITET

BILDREPRODUKTION OCH BILDKVALITÉT

TNM097

Reproduktion av bilder med hjälp av bilddatabas

Viktor SANDBERG

Jacob NYMAN

19 februari 2020



Innehåll

Figurer	2
1 Introduktion	2
2 Metod	2
2.1 Kvalitetsmått	2
2.2 Databas	3
2.3 Bildreproduktion med delta-E	5
2.4 Bildreproduktion med anpassad databas	5
2.5 Bildreproduktion med HVS och delta-E	6
2.6 Bildreproduktion utifrån struktur	6
3 Resultat	7
3.1 Resultat av Bildreproduktion med delta-E	8
3.2 Resultat av Bildreproduktion med anpassad databas	10
3.3 Resultat av Bildreproduktion med HVS och delta-E	11
3.4 Resultat av Bildreproduktion utifrån struktur	11
4 Diskussion	12
5 Slutsats	14

1 Introduktion

Sedan introduktionen av den första iPhone 2007 och sociala mediers framgång har allt fler bilder tagits och delats. Antalet bilder som laddas upp dagligen på sociala medier är miljontals och alla dessa bilder analyseras på ett eller annat sätt. Vikten av att kunna analysera bilder har därför blivit allt viktigare. För att förstå hur analys av bilder fungerar på ett djupare plan har detta projekt handlat om reproduktion av bilder. I detta projekt har bilder reproducerats med hjälp av en bilddatabas för att förstå vad det är som utgör kvalitén i en bild. Är det färgskillnaden? Är det strukturen i bild? Detta är bara några frågor som kommer besvaras i denna rapport. Genom att förstå kvalitén av en bild kommer den också kunna återskapas på ett bättre sätt.

2 Metod

I detta projektet har fyra olika metoder för att reproducera bilder använts. Varje metod med både fördelar och nackdelar som kommer beskrivas i detta kapitel. För att avgöra om en reproduktion har uppnått ett bra resultat används kvalitetsmått ΔE och SSIM, dessa beskrivs mer ingående i kapitel 2.1.

2.1 Kvalitetsmått

För att objektivt kunna avgöra om en reproduktion har gjorts bra, har ekvation 1 och ekvation 2 använts.

$$\Delta E_{ab}^* = \sqrt{(L_2^* - L_1^*)^2 + (A_2^* - A_1^*)^2 + (B_2^* - B_1^*)^2} \quad (1)$$

Ekvation 1 fungerar genom jämföra det euklidiska avståndet (färgskillnaden) mellan två punkter i den enhetsoberoende färgrymden CIELAB. Detta betyder att om ΔE får ett högt värde är avståndet mellan de två punkterna stort, och färgskillnaden är stor [2].

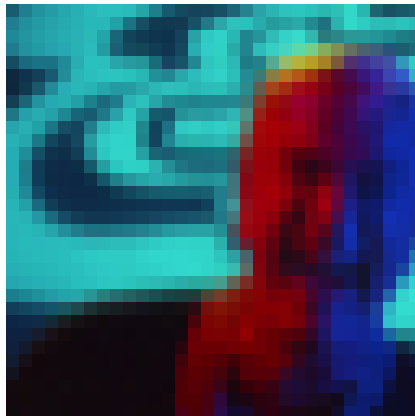
$$\text{SSIM}(x, y) = \frac{(2\sigma_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \quad (2)$$

Ekvation 2 skiljer sig mot ekvation 1 genom att den räknar ut kvalitén av reproduktionen genom att titta på strukturen i bilden. Detta görs genom att

mäta luminansen, kontrasten och strukturen mellan originalbilden x och den återskapade bilden y . I ekvationen motsvarar μ medelvärdet, σ variansen, σ_{xy} kovariansen, samt är c_1 och c_2 variabler för att stabilisera divisionen [1]. Resultatet som kommer ut ur ekvationen ger ett tal mellan $[0,1]$ som avgör hur lik strukturen i reproduktionen är mot originalbilden, ett större tal innebär bättre reproduktion.

2.2 Databas

För att kunna återskapa en bild behövdes också en databas av bilder som kunde utgöra sektioner av bilden som skulle återskapas. I detta projektet skapades bilddatabasen med skärmdumpar från filmer. Valet av bilder till databasen landade på filmer främst av intresse snarare än att det var det bästa valet. I bild 1 kan en bild från bilddatabasen ses.



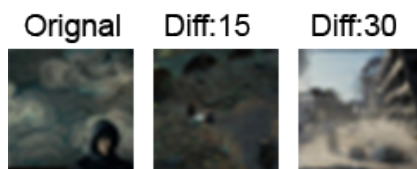
Figur 1: Bild från bilddatabasen i storlek 32x32.

Första delen av skapandet av bilddatabasen var att ladda ner och skala ner bilderna till 32x32 pixlar. Syftet med detta var att reproduceringsbilden inte skulle bli för stor. Något som tänktes på vid nedladdandet av bilder var att få en så stor färg omfång som möjligt. När en stor mängd bilder laddats ner, exakt 1380 bilder skulle databasen optimeras. Anledningen till att databasen skulle optimeras var för att ta bort bilder som låg för nära varandra i färgrymden. För att optimera databasen användes måttet ΔE som presenterades i delkapitel 2.1. För att detta mått skulle gå att applicera var alla bilder i databasen tvungna att göras om från färgrymden RGB till CI-ELAB. Nu fanns ett mått att jämföra skillnaden mellan bilder, men att enbart jämföra skillnaden mellan alla bilder gav inte ett bra resultat. Anledningen till att det resultatet inte blev bra var för att alla bilder har olika

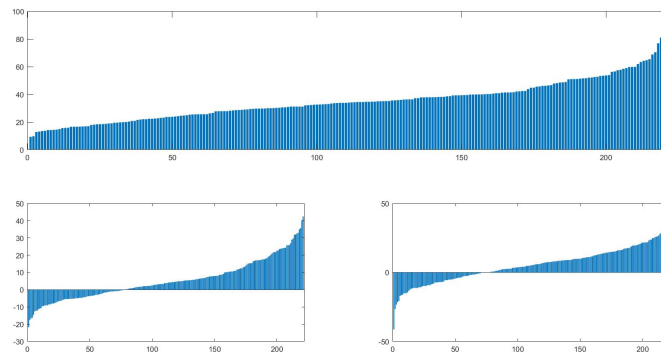
färgskillnad med varandra. För att undvika detta problemet och samtidigt maximera färgrymden i databasen användes ett statussystem. Statussystemet gick ut på att samma bild inte skulle jämföras mer än en gång och så att inte samma bild läggs in i den optimerade databasen flera gånger.

I optimeringen av databasen valdes också ett tröskelvärde för skillnaden mellan bilderna. Tröskelvärdet sattes mellan 15-40 beroende på hur mycket databasen skulle optimeras. I Figur 2 ses en jämförelse mellan olika tröskelvärden. 15 var en liten skillnad mellan bilderna och färgskillanden var i princip inte märkbar vid en skillnad på 15 vilket kan ses i Figur 2. Med det tröskelvärdet resulterade det i att databasen optimerades till storleken 1049 bilder. Målet med optimering var att få ner antalet bilder till 200 eller mindre så ett högre tröskelvärde valdes. När istället ett tröskelvärde på 30 valdes till optimeringen blev storleken istället 127 bilder.

Skillnaden i storleken mellan databaserna blev alltså avsevärt större med det högre tröskelvärdet. Om tröskelvärdet höjs ytterligare kommer databasen bli ännu mindre, detta medför att många färger inte kommer kunna representeras korrekt då databasen blir väldigt begränsad. Ett gränssnitt för att utvärdera färgrymden i databasen utvecklades och kan ses i Figur 3. Med hjälp av detta gränssnitt gick det att se vilka färger databasen hade svårt att representera.



Figur 2: Visar skillnaden mellan en färgskillnad på 15 och 30 i CIELAB.



Figur 3: Visar bilddatabasens färgrymd i CIELAB. Den övre representerar L, de nedre graferna representerar a respektive b.

2.3 Bildreproduktion med delta-E

Den första metoden för att reproducera bilder som implementerades utgick från ekvation 1. Detta gjordes först genom att välja en faktor att skala ner originalbilden med. Detta gjordes för att den återskapade bilden inte skulle bli för stor. Efter detta jämfördes varje pixel i den förminskade versionen av originalbilden, mot alla bilder i databasen. För att en bilddatabasbild skulle bli vald för en viss pixel, jämfördes avståndet mellan pixelns LABmedelvärde mot LABmedelvärdet för varje bild i databasen. Den databasbild som hade den minsta skillnaden mot den jämförda pixelen, valdes att reproducera den specifika pixeln. Den processen kördes till alla pixlar i den förminskade originalbilden hade blivit jämförda. Detta gjorde att varje pixel i den förminskade originalbilden nu representerades med en bild från bilddatabasen.

2.4 Bildreproduktion med anpassad databas

Den andra metodens syfte var att förminska bilddatabasen för reproduktionen ännu mera. För att en bild skulle få vara en del av den nedskalade databasen, jämfördes originalbildens medelvärde i L-kanalen mot medelvärdet i luminansen hos varje bild i bilddatabasen. Om distansen mellan de två värdena var mindre än ett förbestämt tröskelvärde, fick bilden vara med.

Efter att databasen hade blivit förminskad ännu en gång, utfördes samma metod som beskrivit i delkapitel 2.3, med skillnaden att valet mellan bilder som går att välja för att representera en pixel var markant mindre.

2.5 Bildreproduktion med HVS och delta-E

Den tredje metoden som användes vid bildreproduktion var en metod som använde sig av den första metoden 2.3 och *Human visual system* (HVS) [3]. *HVS* är en metod som optimerar en bild eller en del av bild för med avseende på det hur människor ser. På detta sätt upplevs bilden bättre för oss människor. Implementationen av *HVS* som användes i denna metod var samma som användes under Lab 3 i kursen TNM097.

Denna metod fungerade som första metoden beskriven i delkapitel 2.3 med att den letade efter bilden med lägst ΔE . Det denna metoden gjorde annorlunda var att gjorde en extra jämförelse. Denna jämförelsen gjordes när två bilder i databasen hade ΔE värden som var tillräckligt nära varandra efter att de passerat ett tröskelvärde. Jämförelsen gick ut på att applicera *HVS* på båda bilderna och sedan beräkna ΔE värdet och jämföra vilket värde som låg närmast referensen. Anledningen till att denna metod implementerades var för att ta bättre beslut kring databasbilder som var lika varandra och på så sätt välja den bästa bilden utifrån hur människor ser och inte utifrån det lägsta ΔE -värdet.

2.6 Bildreproduktion utifrån struktur

Den sista metoden som implementerades utgick från ett lite annorlunda tillvägagångssätt gentemot de tidigare förklarade metoderna för att återskapa en bild. Syftet med denna metoden var inte att göra en så bra återskapelse som möjligt, utan snarare att försöka fånga strukturen i bilden med så få bilder som möjligt.

Detta gjordes genom att först skala om originalbilden till att vara delbar med storleken på bilderna i databasen. Exempelvis kunde detta innebära att eftersom bilderna i bilddatabasen hade dimensionerna 32x32 pixlar, var en storlek som kunde användes på originalbilden 10 gånger större i varje dimension. Detta innebar att originalbilden skalades till 320x320 pixlar, som skulle reproduceras med 10 stycken databasbilder i varje led.

Efter att bilden hade blivit skalad, jämfördes medelvärde på varje 32x32 yta i bilden mot alla bilder i databasen. De bilder som låg nära i färgrymden mot medelvärde på den jämförda ytan, sparades ner i en ny matris. Dessa bilder jämfördes sedan mot varandra enligt Ekvation 3 där båda kvalitetsmått ΔE och SSIM beskrivna i delkapitel 2.1 användes.

$$\alpha \Delta E + (1 - \alpha) \cdot (1 - \text{SSIM}) \quad (3)$$

Där α i ekvationen är ett tal mellan $[0,1]$, vilket anger hur många procent av metoderna ΔE och SSIM som kommer användas. Ett högt α låter färgreproduktionen i bilden ha mer inflytande, och ett lågt α återger mer struktur i bilden.

Ekvation 3 körs för alla de bilder som låg nära i färgrymden från tidigare steg, och den databasbild som har minsta värdet, är den som är den bästa representanten för den jämförda 32x32 pixelytan.

3 Resultat

Resultatet av detta projektet blev ett antal metoder som gick att använda för att reproducera godtyckliga bilder i alla storlekar. Nedan i Tabell 1, visas hur lång tid varje metod tog att köra och hur väl den reproducerade bilden (se Figur 4) utifrån de två objektiva måtten, SSIM och ΔE .

Tabell 1: Tid och objektiva kvalitetsmått för varje metod gjord se Figur 4.

Name	Time (scale 1/8)	SSIM	ΔE
Metod 2.3	1.9012s	0.1932	24.7993
Metod 2.4	1.031s	0.0155	48.4270
Metod 2.5	172.34s	0.1946	25.5313
Metod 2.6	330.61s	0.2380	22.7696

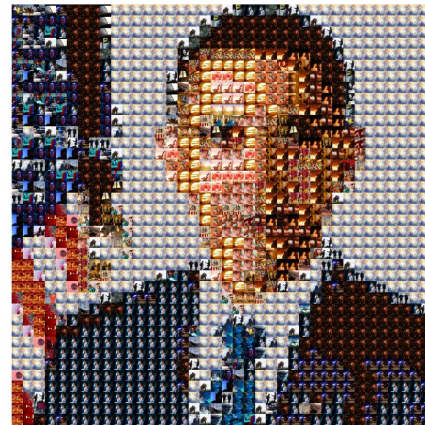
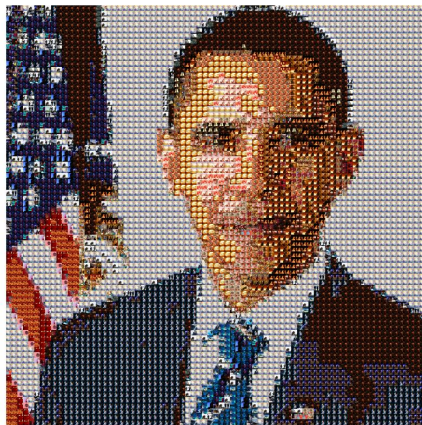
Bilden som skulle återskapas med bilder från bilddatabasen var en bild på den föredetta amerikanska presidenten, Barack Obama, se Figur 4. Denna bild skalades ner 8, respektive 4 gånger, i varje metod för att ge jämförbara resultat mellan de olika metoderna. Bildreproduktionerna för varje metod kan ses i delkapitel 3.1, 3.2, 3.3 och 3.4.



Figur 4: Originalbilden för bildreproduktion

3.1 Resultat av Bildreproduktion med delta-E

Figur 5 visar de två reproduktionerna som använde sig av metoden diskuterad i delkapitel 2.3. Figur 5a visar reproduktionen av originalbilden när den blivit nedskalad till en fjärdedel av originalstorleken och Figur 5b visar när den blivit skalad en åttondel.



(a) Skala: $1/4$ ΔE : 20.2385 SSIM: 0.2721 (b) Skala: $1/8$ ΔE : 24.7993 SSIM: 0.1932

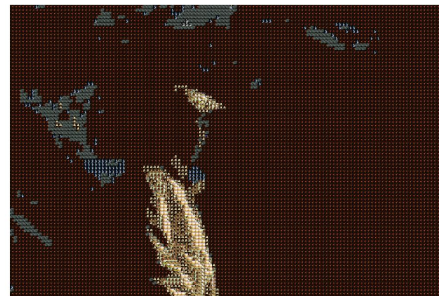
Figur 5: Resultatet av reproduktionen av bilden på Obama med metod 2.3

Förutom bilden på Obama gjordes även reproduktion på en mörk bild som kan ses i Figur 6, en överexponerad i Figur 7 och en landskapsbild i Figur 8.

I detta fallet har alla bilder blivit nedskalade till en åttondel.



(a) Mörk bild

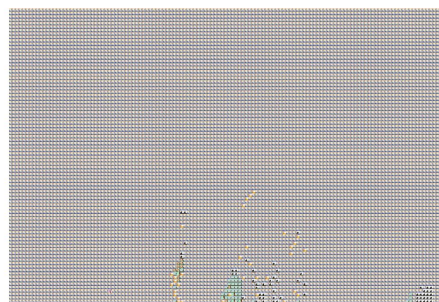


(b) Skala: 1/8 ΔE : 16.8917 SSIM: 0.1435

Figur 6: Resultatet av reproduktionen av den mörka bilden med metod 2.3



(a) Överexponerad bild

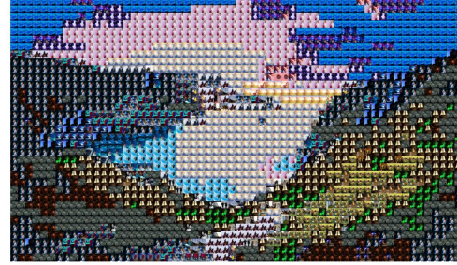


(b) Skala: 1/8 ΔE : 29.9339 SSIM: 0.0429

Figur 7: Resultatet av reproduktionen av den överexponerad bilden med metod 2.3



(a) Landskapsfotografi

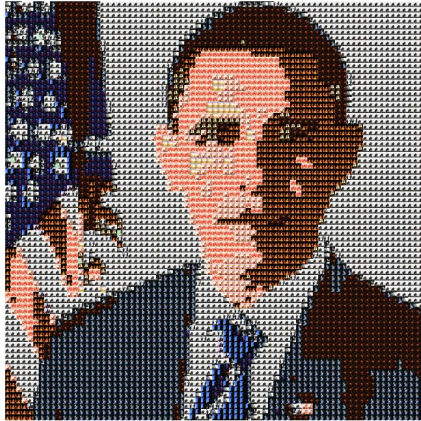


(b) Skala: 1/8 ΔE : 26.9279 SSIM: 0.1757

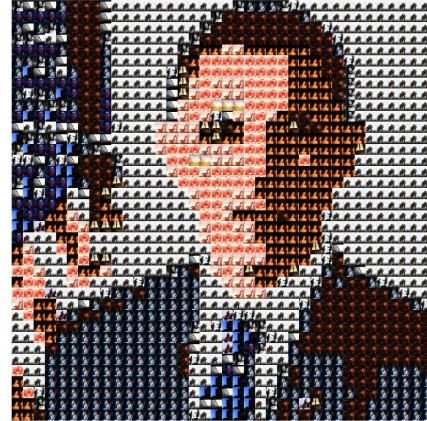
Figur 8: Resultatet av reproduktionen av landskapsbilden med metod 2.3

3.2 Resultat av Bildreproduktion med anpassad databas

Figur 9 visar de två reproduktionerna som använde sig av metoden diskuterad i delkapitel 2.4. Figur 9a visar reproduktionen av originalbilden när den blivit nedskalad till en fjärdedel av originalstorleken och Figur 9b visar när den blivit skalad en åttondel.



(a) Skala: 1/4 ΔE : 24.0170 SSIM: 0.2055

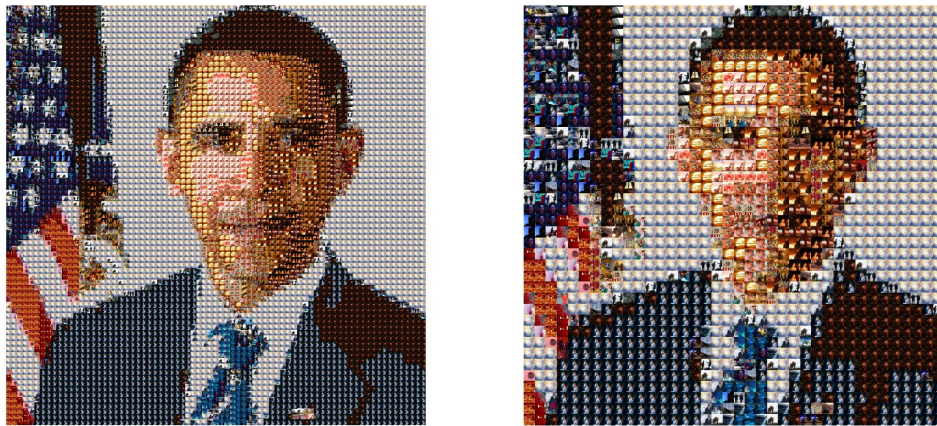


(b) Skala: 1/8 ΔE : 48.4131 SSIM: 0.032

Figur 9: Resultatet av reproduktionen av bilden på Obama med metod 2.4

3.3 Resultat av Bildreproduktion med HVS och delta-E

Figur 10 visar de två reproduktionerna som använde sig av metoden diskuterad i delkapitel 2.5. Figur 10a visar reproduktionen av originalbilden när den blivit nedskalad till en fjärdedel av originalstorleken och Figur 10b visar när den blivit skalad en åttondel.

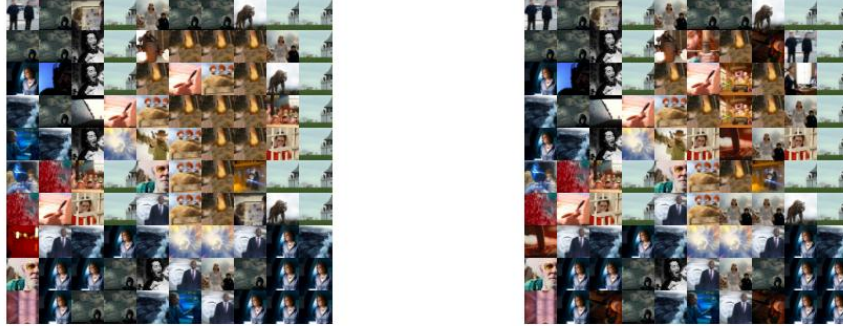


(a) Skala: 1/4 ΔE : 19.8477 SSIM: 0.2785 (b) Skala: 1/8 ΔE : 25.5313 SSIM: 0.1936

Figur 10: Resultatet av reproduktionen av bilden på Obama med metod 2.5

3.4 Resultat av Bildreproduktion utifrån struktur

Figur 11 visar de två reproduktionerna som använde sig av metoden diskuterad i delkapitel 2.6. Figur 11a visar reproduktionen av originalbilden när $\alpha = 0.2$ och Figur 11b visar när $\alpha = 0.8$



(a) $\alpha = 0.2\Delta E$: 28.1618 SSIM: 0.2353

(b) $\alpha = 0.8\Delta E$: 28.3742 SSIM: 0.2268

Figur 11: Resultatet av reproduktionen av bilden på Obama med metod 2.6

4 Diskussion

Resultatet av bildreproduktion var av olika kvalité beroende på vilken metod som användes. Den metod som gick snabbast att köra var metod 2.4. De två metoderna som var mer komplexa 2.5 och 2.6 var istället minuter långsammare vilket också visas i resultatet. Frågan är dock om resultatet av dessa komplexa metoder blev bättre? Enligt de objektiva kvalitetsmåten var metod 2.4 den sämsta. Den bästa metoden enligt de objektiva kvalitetsmåten var metod 2.6 den bästa, men inte långt före metod 2.3. Den bästa metoden om tid och objektivkvalité kombineras är 2.3. Metod 2.5 tog närmare 3 minuter att köra och hade in princip likadana värden som metod 2.3 vilket inte gör att *HVS*-metoden inte är värd att använda. Även om metod 2.4 hade sämst värden enligt de objektiva kvalitetsmåten, skulle den med största sannolikhet få ett bättre resultat om ett kvalitetsmått med *HVS* hade använts, för den reproducerade bild fångar väsentliga i originalbilden.

Den överexponerade bilden Figur 7 och den mörka bilden Figur 6 som reproducerades med metod 2.3 gav ett dåligt resultat. Dåligt i form av att metoden inte kunde reproducera de ljusa och mörka partierna i bilden utan endast reproducerades homogen ytor. För att reproducera ett bättre resultat hade en annan metod behövt användas. En metod som hade kunnat återskapa ett bättre resultat är metod 2.6. Denna metod hade reproducerat ett bättre re-

sultat då den inte enbart hade tittat på färgskillnaden i bilden utan också strukturen.

Alla metoderna användes inte i reproduktionen vid olika typer av bilder. Detta är dock något som borde gjorts då metoderna med största sannolikhet är bra på att reproducera olika typer av bilder.

Vidare arbete med de metoder som utvecklats är att optimera de mer och att göra de mer komplexa metoderna mer tidseffektiva. Särskilt metod 2.5 då delar av de operationer som gjordes i den kunde gjorts i förväg. En annan metod som kunde använts hade varit att rotera bilderna för att ännu bättre stämma överens med originalbilden. Alla metoderna som utvecklats bygger på att beräkna ΔE vilket är färgskillnaden mellan bilderna. Detta är dock ett medelvärde över bilden från databasen och området i originalbilden. För att skapa en bättre reproduktion måste flera detaljer tas i beräkning. Ett exempel på detta är om en bild i databasen består av två dominanta färger vars medelvärde inte ger bilden rättvisa. För att kunna använda sådana bilder på ett optimalt sätt behövs en annan metod.

I ett tidigt skede av projektet experimenterades det med flera olika metoder för att optimera bilddatabasen. Den första metoden som experimenterades med var att beräkna medelvärdet för varje RGB kanal av en bild. Detta gav dock inget bra resultat och det liknade inte alls färgen av originalbilden. När den metoden var bortsedd från började arbete med en mer ambitiös metod: *k-means clustering*[4]. Den metoden utgick från att hitta den mest dominanta färgen i en bild och spara dessa i kluster. Då metoden låter användaren själv bestämma antalet kluster blev detta svårt att applicera på bilder av olika karaktär. En bild kan ha två dominanta färger och då är två ett bra antal kluster, men om bild är uppbyggd av många färger blir inte två kluster ett bra val och då blir det istället att ett medelvärde måste tas inom ett specifikt kluster. På grund av denna anledning avfärdades även den metoden för att bestämma färgen i en bild. Efter detta började ΔE användas för att skapa en databas med en optimal färgrymd.

Om *k-means clustering* hade använts i kombination med ΔE hade förmodligen både uppbyggnaden av databasen samt jämförelsen mot originalbilden kunnat göras bättre. *k-means clustering* är dock en algoritm som är relativt tidskrävande och frågan är om den extra tiden det skulle ta är värt det för ett marginellt bättre resultat. Från de resultaten som producerats går det att se att skillnaden mellan de mer avancerade metoderna och den enklare som enbart använde sig av ΔE är marginell även om tidsåtgången för att skapa dem är flera minuter längre.

Innan databasen optimerats innehöll den 1400 bilder från filmer. Ett problem med bilder från filmer är att majoriteten av bilderna innehåller en människa som ständigt påverkar färgen i bilden. Många av bilderna i databasen var också väldigt lika i att de var mörka eller var ett porträtt på en människa som påverkade färgen. Detta märktes speciellt när databasen optimerades då den gick från 1400 till 127 när ett tröskelvärde på färgskillnaden var 30, se Figur 2. Databasens färgrymd kan ses i Figuren 3 och där ses det tydligt att vissa färger inte blivit representerade. Detta problemet kunde motarbetats på flera sätt. En större databas hade lett till ett större urval av bilder som med största sannolikhet hade kunnat representera en större färgrymd. Problemet kunde även ha motarbetats genom att använda sig av bättre bilder som inte främst bestod av människor.

5 Slutsats

Resultatet från detta projekt har givet en ny syn på bildreproduktion och bildkvalitet. För att reproducera en bild på bästa sätt räcker det inte att ha en metod. Då olika metoder fungerar bättre på olika typer av bilder. Skillnaden i tid mellan olika metoder är också markant vilket kan ses i Tabell 1. Skillnaden i tid för att få ett marginellt bättre resultat är något som framgått i detta projekt. För att kunna reproducera en bild som tar hänsyn till många faktorer kostar det priset av längre tid. Det har framgått att det finns otaligt många sätt att reproducera bilder och att det går att göra mycket mer komplext än vad som har gjorts i detta projektet. Samma sak gäller skapandet och optimerandet av en databas, det kan göra på otalig många sätt. Slutligen från detta projektet har det framgått hur mycket som kan göras och vidareutvecklas snarare än motsattsen. Detta projektet har öppnat dörren till vad som faktiskt är möjligt inom bildreproduktion och bildkvalitet.

Referenser

- [1] B. Kumar, S. B. Kumar, and C. Kumar. Development of improved ssim quality index for compressed medical images. In *2013 IEEE Second International Conference on Image Information Processing (ICIIP-2013)*, pages 251–255, Dec 2013.

- [2] L. Mandic, S. Grgic, and M. Grgic. Comparison of color difference equations. In *Proceedings ELMAR 2006*, pages 107–110, June 2006.
- [3] B. Wang, Z. Wang, and and. Hvs-based structural similarity for image quality assessment. In *2008 9th International Conference on Signal Processing*, pages 1194–1197, Oct 2008.
- [4] Z. Yong and X. Shixiong. A novel clustering algorithm based on hierarchical and k-means clustering. In *2007 Chinese Control Conference*, pages 605–609, July 2006.