

Meetrapport Vision

Maaïke Hovenkamp

1741256

Duur Alblas

1739724

11-03-2020

| | |
|-------------------|----------|
| Doel | 2 |
| Hypothese | 2 |
| Werkwijze | 2 |
| Resultaten | 3 |
| Verwerking | 4 |
| Conclusie | 5 |
| Evaluatie | 5 |

Doel

Welk van de 5 gray scaling methoden geeft ons het meest accurate facial recognition resultaat, dus herkend het vaakst een menselijk gezicht als wij dat ook doen en niet als wij ook geen gezicht herkennen, zonder er significant langer over te doen en zonder de andere processen van de pipeline aan te passen?¹

Dit willen wij uitzoeken voor gebruik voor het opstellen van een profiel van misdadigers in een misdadigers database. Het opstellen van dit profiel vereist een monochrome afbeelding.

Door dit onderzoek te doen kunnen we bij de conclusie aantonen welke methode het beste werkt voor onze onderzoeksvraag aan de hand van onze resultaten.

Hypothese

Onze hypothese is dat Luma of Intensity ons het beste resultaat zal geven op onze onderzoeksvraag. Wij denken Luma omdat hij gamma correctie toepast en het resultaat dus beter herkenbaar zou moeten zijn voor het menselijk oog. Ook denken wij dat Intensity een goed resultaat kan leveren omdat het als een simpele, maar passende methode wordt gezien.

Werkwijze

We gaan onze metingen uitvoeren door middel van de methoden op 7 sets van afbeeldingen uit te voeren met behulp van Arno Kamphuis' pipeline.

1 van de sets die we gebruiken bevat de afbeeldingen die in de repository van Arno Kamphuis zitten.

Deze set is aanzienlijk kleiner dan de andere sets en bevat maar 6 afbeeldingen van gezichten.

De andere 6 sets hebben 36 afbeeldingen per set, we hebben afbeeldingen van gezichten, natuur, kunst en dieren. Deze andere afbeeldingen zijn copyrighted en zullen en mogen dus niet publiek gemaakt worden.

Nadat we alle methoden op de sets hebben toegepast gaan wij met de hand noteren welke afbeeldingen wel of niet accuraat herkend zijn. Door de main.cpp van het framework van Arno aan te passen worden alle afbeeldingen automatisch door een methode gehaald en opgeslagen. Met behulp van deze data gaan wij aantonen hoe accuraat elke methode is.

¹ Getest en geïmplementeerd op de pipeline van Arno Kamphuis, 15-02-19, HU-Vision-1819-Base, <https://github.com/arnokamphuis/HU-Vision-1819-Base>

Resultaten

Hier volgen de resultaten van het uitgevoerde experiment. De tabellen zijn verdeeld onder de verschillende testsets en laten zien in hoeveel van de afbeelding accuraat een gezicht wel of niet herkend is. Set A bevat de afbeeldingen van Arno Kamphuis. Het woord Arno staat voor de standaard geïmplementeerde methode van Arno kamphuis. Sommige afbeeldingen geven een error op wanneer het programma gedraaid word. Deze afbeeldingen verschillen per methode of ze zorgen voor een error of niet. Omdat er verschil in zit hebben de afbeeldingen die regelmatig een error opleveren niet verwijderd maar in onze resultaten opgenomen. Veel voorkomende afbeeldingen met problemen waren Man en Teenager.

| | Random A | | | Random B | | | Random C | | |
|-----------|----------|--------------|-------|----------|--------------|-------|----------|--------------|-------|
| | Accurate | Not accurate | error | Accurate | Not accurate | error | Accurate | Not accurate | error |
| Arno | 24 | 11 | 1 | 26 | 9 | 1 | 26 | 9 | 1 |
| Intensity | 27 | 8 | 1 | 27 | 9 | 0 | 27 | 8 | 1 |
| Value | 28 | 7 | 1 | 27 | 8 | 1 | 28 | 7 | 1 |
| Luster | 26 | 7 | 1 | 26 | 10 | 0 | 26 | 9 | 1 |
| Luminance | 26 | 8 | 2 | 26 | 10 | 0 | 26 | 9 | 1 |
| Luma | 27 | 9 | 0 | 26 | 10 | 0 | 26 | 10 | 0 |

| | Random D | | | Random E | | | Random F | | |
|-----------|----------|--------------|-------|----------|--------------|-------|----------|--------------|-------|
| | Accurate | Not accurate | error | Accurate | Not accurate | error | Accurate | Not accurate | error |
| Arno | 26 | 10 | 0 | 26 | 10 | 0 | 26 | 9 | 1 |
| Intensity | 27 | 9 | 0 | 25 | 11 | 0 | 25 | 9 | 2 |
| Value | 25 | 11 | 0 | 24 | 12 | 0 | 25 | 10 | 1 |
| Luster | 27 | 9 | 0 | 25 | 11 | 0 | 26 | 9 | 1 |
| Luminance | 27 | 9 | 0 | 24 | 11 | 1 | 24 | 10 | 2 |
| Luma | 27 | 9 | 0 | 25 | 11 | 0 | 25 | 11 | 0 |

| | Set A | | |
|-----------|----------|--------------|-------|
| | Accurate | Not accurate | error |
| Arno | 5 | 2 | 0 |
| Intensity | 5 | 2 | 0 |
| Value | 2 | 4 | 0 |
| Luster | 2 | 4 | 1 |

| | | | |
|-----------|---|---|---|
| Luminance | 5 | 2 | 0 |
| Luma | 0 | 6 | 0 |

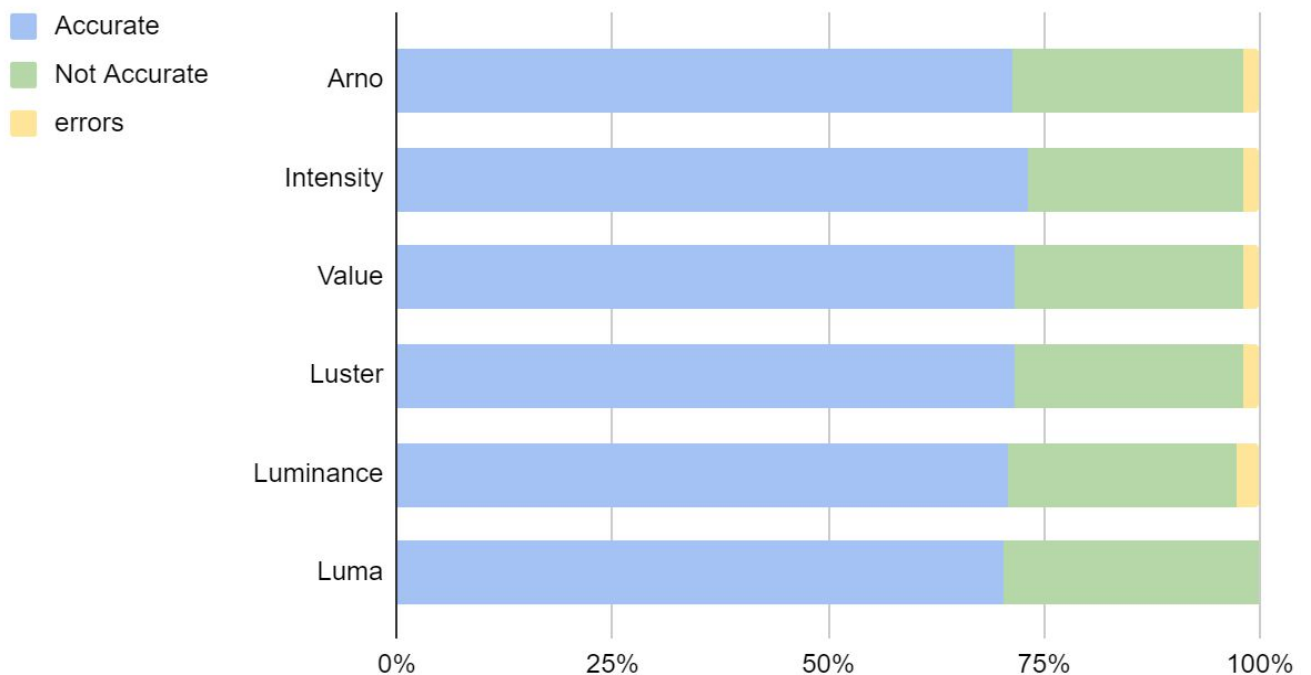
Verwerking

Er wordt eerst een totaal berekend van de verschillende categorieën per methode en deze totalen worden omgezet in percentages.

| | Accurate | Not Accurate | errors | Percentage accurate | Percentage not accurate | Percentage error |
|-----------|----------|--------------|--------|---------------------|-------------------------|------------------|
| Arno | 159 | 60 | 4 | 85,02673797 | 32,0855615 | 2,139037433 |
| Intensity | 163 | 56 | 4 | 87,1657754 | 29,94652406 | 2,139037433 |
| Value | 159 | 59 | 4 | 85,02673797 | 31,55080214 | 2,139037433 |
| Luster | 158 | 59 | 4 | 84,49197861 | 31,55080214 | 2,139037433 |
| Luminance | 158 | 59 | 6 | 84,49197861 | 31,55080214 | 3,20855615 |
| Luma | 156 | 66 | 0 | 83,42245989 | 35,29411765 | 0 |

Hieronder de resultaten uitgezet in een diagram voor snel overzicht. De gevonden verschillen in de implementaties zijn minimaal.

Resultaten meetproef in procenten



Met de resultaten wordt de standaarddeviatie berekend en wordt er een uitspraak gedaan over de accuraatheid van de verschillende methodes.

| | Accurate STDEV | Answers |
|-----------|----------------|-----------------------|
| Arno | 7,846746368 | 85 +- 7.85 σ % |
| Intensity | 8,118174789 | 87 +- 8.12 σ % |
| Value | 9,268482179 | 85 +- 9.27 σ % |
| Luster | 9,089502056 | 84 +- 9.01 σ % |
| Luminance | 7,828519291 | 84 +- 7.83 σ % |
| Luma | 9,860937848 | 83 +- 9.86 σ % |

Uit de resultaten blijkt dat Luma het slechtste presteert, er zijn weliswaar geen errors maar dat is te verklaren doordat er helemaal niets herkend wordt met de Luma methode. Er wordt geen enkel gezicht gevonden, niet waar er wel gezichten zijn en ook niet waar er geen gezichten zijn. Doordat er geen oren worden gevonden kapt het programma zichzelf af waardoor er geen error wordt gevonden. De Intensity methode daarentegen presteert het beste. Deze methode herkende het vaakste accuraat gezichten waar die ook aanwezig waren en vond geen gezichten waar die ook niet waren.

Conclusie

Uit de resultaten kunnen wij concluderen dat de Intensity methode het beste presteert binnen het kader van onze hoofdvraag. Wij hebben deze conclusie getrokken uit de resultaten die uit het experiment gekomen zijn. Intensity heeft het hoogste percentage voor accuraatheid behaald met een 2% voorsprong op het framework van Arno en de methode value. Dit zijn minimale verschillen. Ook bracht intensity niet meer errors met zich mee dan deze twee andere methoden waarmee hij voor de top strijd.

Evaluatie

Tijdens het onderzoek zijn er veel dingen goed gegaan maar er zijn ook dingen die niet goed gingen. Het is ons goed gelukt onze methode te implementeren in Arno's pipeline. Ook zijn we tevreden over onze documenten.

Wat er minder goed ging is dat als snel bleek dat diverse afbeeldingen errors veroorzaakte in het programma van Arno. Dit hebben we opgelost door met onze docent, Diederik Yamamoto-Rooijers, te overleggen. We kwamen tot het besluit dat we ook zouden bijhouden hoeveel afbeeldingen een error gaven. De verschillende methodes gaven in meestal bij dezelfde afbeeldingen een error aan maar het is een of twee keer gebeurd dat er een ander plaatje de oorzaak was. We weten nog steeds niet wat deze errors veroorzaakt.

In het begin ging het werk verdelen goed aangezien er in de image shells een aantal duidelijk verdeelbare functies aanwezig waren. Later werd het lastiger het werk te verdelen aangezien we maar op 1 laptop Visual Studio werkend hadden en een groot deel van het onderzoek bestaat uit het runnen

van het programma met als enige afwisseling dat er af en toe een afbeelding die een error opleverde en verwijderd moest worden. Nadat de resultaten klaar waren kon het werk weer eerlijk verdeeld worden. De afbeeldingen moesten met de hand worden geëvalueerd wat veel tijd kostte. Hier konden we dus de methoden mooi verdelen onder ons tweeën. Voor de volgende keer zou het beter zijn om de software op twee laptops te draaien zodat de run time omlaag kan.