

ИСПОЛЬЗОВАНИЕ ШКАЛ ОТТЕНКОВ БАЗОВЫХ ЦВЕТОВ ДЛЯ РЕШЕНИЯ ЗАДАЧИ БИНАРИЗАЦИИ

1. Постановка задачи

Целью данной работы является анализ алгоритма пороговой бинаризации сложного изображения, исследование его особенностей и преимуществ по сравнению с альтернативными способами бинаризации. Под сложным изображением будем понимать изображение в оттенках серого с низким уровнем контраста, с яркостной неравномерностью фона, присутствием шума. Так, например, результаты обработки изображения спортсмена предполагают облегчить работу судей, сделать процесс судейства автоматизированным, избавиться от доли субъективности, которая присутствует при судействе.

При оценивании выступлений спортсменов в художественной гимнастике судьи принимают решение о получаемой оценке на основе правильности выполненного элемента. Примером может быть элемент, который называется «прыжок шагом». Элемент считается правильно выполненным, если гимнастка раскрывает ноги в воздухе в полный шпагат (угол между ногами составляет 180° или больше). В этом случае гимнастка получает 0,2 балла. Иначе, элемент считается не выполненным и оценка - 0 баллов (рис.1).

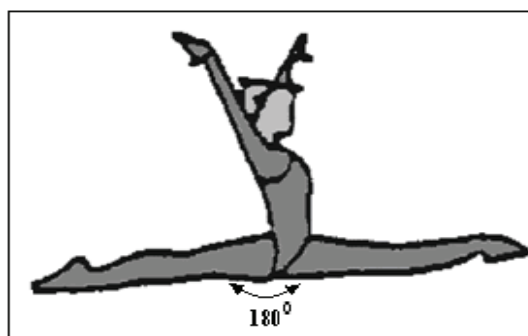


Рисунок 1 – Пример элемента, который подлежит оцениванию

Для того чтобы сделать процесс судейства автоматизированным, предлагается выполнить следующие шаги:

- бинаризация изображения - приведение к черно-белому виду или к изображению, содержащему два контрастных цвета одной цветовой шкалы

© Барковская О.Ю., Аксак Н.Г., 2006

(выполняется одновременно с разбиением изображения на два кластера, что необходимо для дальнейшего распознавания);

- скелетонизация - выделение срединных осей (остовов) на бинарном изображении, используя алгоритм морфологической скелетонизации или случайную обработку;
- распознавание. В данном случае выставление оценки за выполненный элемент.

Необходимость бинаризации заключается в том, что все методы распознавания образов будут хорошо работать лишь тогда, когда они основываются на характеристиках, которые не зависят от конкретных методов получения изображения, т.е. когда можно оценить истинные свойства объекта, например, форму.

Поэтому необходимость сведения решения задачи для цветного изображения к решению задачи для бинарного изображения становится очевидной. Наиболее целесообразно привести изображение к бинарному простыми методами, учитывая яркость/цвет пикселей [3], а далее работать с бинарным изображением. Среди методов получения бинарного изображения существует также и бинаризация на основе нейросетей. Поскольку задача бинаризации является своего рода задачей классификации и сегментации, то она без труда решается нейронными сетями. Бинаризация на основе нейронных сетей является альтернативой другим методам получения бинарного изображения.

В статье предлагается способ бинаризации, основанный на нахождении порога яркости изображения, который делит исходное изображение на 2 класса: класс «фон» и класс «основное изображение».

Предположим, что бинаризованное изображение будет состоять из точек двух типов: фоновых точек и точек интереса. Пусть 0 - фон, 1 - основное изображение.

2. Шкалы оттенков первичных и вторичных цветов

Формирование пикселя любого цвета производится согласно трехкомпонентной теории цветового деления, в основе которой лежит возможность представления цвета сложением в нужной пропорции трёх основных (первичных) цветов: красного, зеленого, синего. Именно эти три цвета выбраны за основные, так как каждый из них нельзя получить смешением двух других и смешение которых в

разных соотношениях даёт возможность получить другие цвета. При наложении попарно трех первичных цветов, образуются вторичные цвета: голубой, пурпурный, желтый. Так, например, вторичный желтый цвет образуется путем сложения первичных красного и зеленого, а синий в формировании желтого не участвует. Если объединить красный, зеленый и синий свет, то длины волн суммируются и в результате получится луч белого света. А если объединить первичный синий и вторичный желтый, то так же получится белый свет.

В современных компьютерах воспроизводится 256 (от 0 до 255) значений каждого из трех цветов, следовательно, общее количество возможных цветов на мониторе: $\sum C = 256^3 = 16700000$, где C – цвет пикселя.

Определенный цвет может быть описан значением цветности и яркости.

Положим, что пиксель исходного изображения $f(x,y)$, где x,y – координаты пикселя, обладает определенным цветом C . Составными частями цвета C являются основные цвета R, G, B со значением интенсивности составляющих $R(C), G(C), B(C)$. Тогда яркость $L(C)$ цвета C определяется по следующей формуле:

$$L(C) = n_{21}R(C) + n_{22}G(C) + n_{23}B(C), \quad (1)$$

где n_{21}, n_{22}, n_{23} – постоянные, которые определяются свойствами человеческого зрения и составляют: $n_{21} = 0,3$, $n_{22} = 0,59$, $n_{23} = 0,11$.

Значение яркости $L(C)$ лежит в диапазоне $0 \leq L(C) \leq 255$. Это вытекает из того, что переменные $R(C), G(C), B(C)$ принимают значения от 0 до 255.

2.1 Формирование шкал оттенков основных цветов

Формирование цвета в RGB-модели можно представить в виде трехмерной системы координат (рис.2) [1].

Оси координат x, y, z соответствуют интенсивности составляющих красного, зелёного и синего цветов в RGB-модели и имеют максимальное значение координат 255. Когда интенсивность всех трех составляющих равна нулю $(R,G,B)=(0,0,0)$ мы получаем чёрный цвет. Он соответствует началу системы координат. Белый цвет формируется при сложении всех цветов максимальной

интенсивности, т.е. $(R,G,B)=(255,255,255)$, что также показано на рис.2.

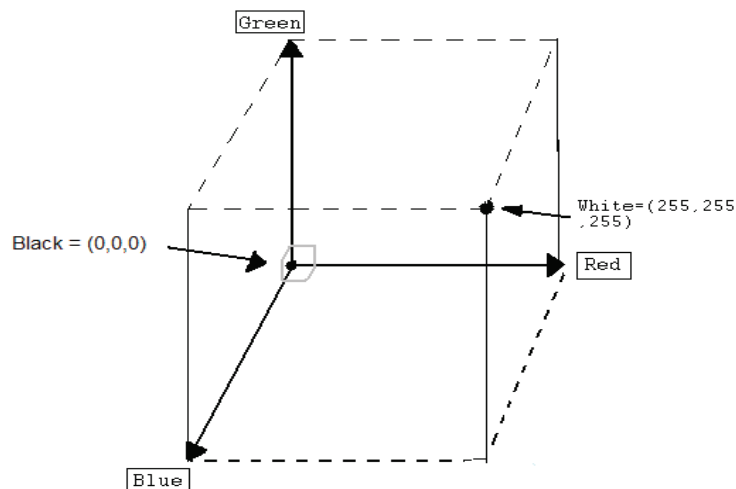


Рисунок 2 – Представление цвета в трехмерном пространстве

Для нахождения порога яркости изображения спортсменки, используемого при распознавании выполненного элемента, удобно использовать изображения в оттенках одного конкретного цвета. Наиболее часто используют шкалы оттенков первичных и вторичных цветов трехкомпонентной теории цветового деления: оттенки красного, синего, зеленого, голубого, желтого и розового. Данные цвета являются вершинами куба, описанного на рис.2. Шкала яркостей полутонового изображения в оттенках любого из указанных цветов включает в себя цвета от чёрного ($R=G=B=0$) до белого ($R=G=B=255$) с конкретным центральным цветом шкалы. Если, например, центральным цветом шкалы является красный цвет $(R,G,B)=(255,0,0)$, то данное изображение представлено в оттенках красного и т.д. Можно выделить также шкалу оттенков серого. Ей надо уделить особое внимание, поскольку серый цвет не является ни первичным, ни вторичным для RGB-модели и не является вершиной куба.

Рассмотрим диапазон цветов, которые относятся к оттенкам красного цвета. Шкала оттенков красного показана на рис.3.

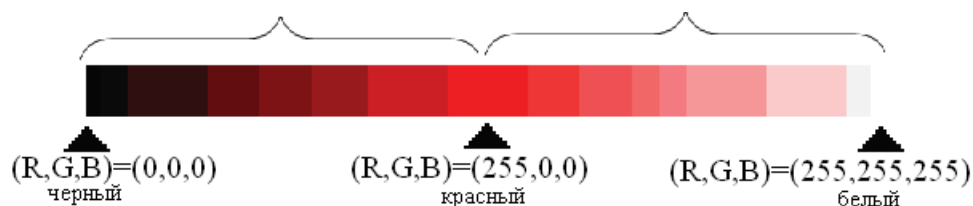


Рисунок 3 – Шкала цветов оттенков красного цвета

Очевидно, что в формировании красного цвета принимает участие лишь составляющая красного цвета, а интенсивности синего и зеленого равны нулю, т.е. $(R,G,B)=(255,0,0)$. При уменьшении яркости красного цвета, интенсивности синего и зеленого остаются равными нулю, а интенсивность красной составляющей постепенно приближается к нулю до тех пор, пока не будет достигнуто значение $(R,G,B)=(0,0,0)$. При увеличении яркости красного, интенсивность красной составляющей остаётся равной 255, а интенсивность синей и зелёной составляющих постепенно увеличиваются до значения $(R,G,B)=(255,255,255)$. Т.е. всю шкалу оттенков красного (как и шкалы оттенков остальных цветов) логически можно поделить на две равные части:

1. от черного до красного: $(0,0,0) \leq (R,G,B) \leq (255,0,0)$;
2. от красного до белого: $(255,0,0) \leq (R,G,B) \leq (255,255,255)$.

Граничным (пороговым) значением является красный цвет. При выполнении бинаризации (рассматриваются только два цвета – черный и белый) левая часть шкалы становится черной, а правая – белой (рис.4).



Рисунок 4 – Бинаризованная шкала оттенков цвета

В векторном виде это можно представить:



Рисунок 5 – Векторное представление бинаризованной шкалы

Для всех семи шкал оттенков первичных и вторичных цветов RGB-модели бинаризация выполняется по принципу, описанному выше, меняются только пороговые значения. Поскольку все первичные и вторичные цвета являются вершинами куба (рис.2), то и весь диапазон цветов, которые относятся к оттенкам конкретного цвета можно определить по этому кубу. Вернёмся к примеру с оттенками красного цвета. Всю шкалу оттенков красного можно разбить на две части: черный – красный, красный – белый. На рис.6 такое бинарное разбиение шкалы показано двумя векторами: \vec{GA}

\vec{AH} , где начало координат (интенсивность всех составляющих минимальна) соответствует белому цвету, а вершина с координатами (255,255,255) чёрному. Все значения существующих оттенков красного, так же, как и оттенков остальных цветов, лежат на ребрах куба, которые являются векторами, описывающими шкалы оттенков.

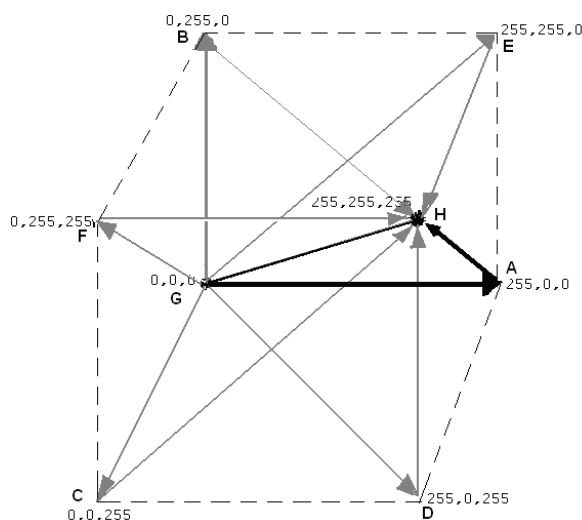


Рисунок 6 – Шкалы яркостей оттенков цветов

Наиболее удобной шкалой для выполнения бинаризации является шкала оттенков серого. Это означает, что цветовая шкала включает в себя цвета от чёрного ($R=G=B=0$) до белого ($R=G=B=255$) с центральным цветом шкалы серым ($R=G=B=128$) (рис.7). Для изображений данного типа значения всех трех составляющих равны, т.е. полученные значения яркости копируются во все три канала $R=G=B=L(C)$. Поскольку серый цвет не является вершиной куба, то весь диапазон цветов оттенков серого возможно представить лишь 1 вектором, соединяющим начало координат (черный цвет) с вершиной, соответствующей белому цвету. Все оттенки серого лежат на этом векторе. Для выполнения бинаризации выбирается порог, который лежит посередине вектора, т.е. цвет с координатами $(R,G,B)=(128,128,128)$. Бинаризованная шкала показана на рис.4.



Рисунок 7 - Шкала яркостей полутонового изображения в оттенках серого

3. Способ выполнения бинаризации на основе порога

Метод пороговой бинаризации [2] для преобразования изображения в градациях серого в черно-белое изображение базируется на получении порога яркости изображения. Под *порогом яркости* будем понимать значение, относительно которого выполняется разделение на черные и белые области изображения.

Способ бинаризации на основе порога.

1. Определение размера изображения $M \times N$ пикселей (размер пикселя компьютерного монитора колеблется от 0,2 до 0,35 мм) и вычисление общего количества пикселей;

2. Определение уровня яркости каждого пикселя исходного изображения $f(x,y)$;

3. Построение гистограммы (графика, показывающего частоту попаданий значений переменной в отдельные интервалы) по полученным значениям яркостей изображения. Гистограмма строится по следующему правилу: по оси абсцисс указывается возможный наибольший интервал изменения яркости от 0 до 255. По оси ординат указывается число пикселей K с близким значением уровня яркости f , которые попадают в интервал от $T-\Delta T$ до $T+\Delta T$, где T - предполагаемое пороговое значение.

При использовании шкалы полутонов возможно построение четырех различных гистограмм, по которым будет найден порог. Это гистограммы, построенные по трём каналам – интенсивности трёх составляющих любого цвета (красный, зеленый, синий), и по рассчитанному значению яркости пикселя. В случае изображений в оттенках красного, синего, зеленого, голубого, желтого и розового цветов все 4 гистограммы будут разными для каждого оттенка. Но, гистограммы, построенные по изображению в градациях серого, будут одинаковы, т.к. значения интенсивности каждой составляющей цвета и значение яркостей всегда одинаковы: $R=G=B=L(C)$. Таким образом, достаточно построить лишь одну гистограмму для всего изображения в градациях серого, которая будет соответствовать гистограммам по каждому каналу и по яркости, и по ней находить порог бинаризации.

Это является большим преимуществом использования оттенков серого для распознавания образов и облегчает задачу нахождения порога при бинаризации изображения;

4. На основе анализа гистограммы выбираем порог. Наиболее просто выбрать порог в случае бимодального распределения, полученного в результате качественной оценки характеристик распределения гистограммы. Если получено бимодальное распределение, т.е. имеющее 2 пика, то разделения выборки яркости пикселей на две части не составляет труда. В этом случае мы получаем единственный порог T , который разбивает значения, лежащие на оси абсцисс на два участка с наибольшей интенсивностью: $0 \leq T \leq 255$.

Этот порог не обязательно будет находиться в центре оси абсцисс. Пороговое значение яркости зависит от контрастности изображения и от того, какие оттенки преобладают - темные или светлые. Если в бинаризуемом изображении преобладают темные оттенки, то пороговое значение будет смещено от середины к 0, а в случае преобладания светлых оттенков, наоборот, к 255. Это вытекает из того, что значения уровней яркости полутонового изображения (изображение в градациях серого) распределены следующим образом: значению яркости (0,0,0) соответствует черный цвет, а значению яркости (255,255,255) белый (рис. 7).

После определения порога выполняется сравнение значения яркости каждого пикселя с пороговым значением яркости.

Бинарная классификация

- Если значение яркости бинаризуемого пикселя больше определенного порогового значения яркости T , то пиксель окрашивается в белый цвет (фон);
- Если значение яркости бинаризуемого пикселя меньше определенного порогового значения яркости T , то пиксель окрашивается в черный цвет (образ).

Математическое описание того, что яркость квантуется (принимает значения только из конечного набора величин) представляется в виде системы уравнений (2).

$$g(x,y) = \begin{cases} 1, & \text{если } f(x,y) > T; \\ 0, & \text{если } f(x,y) \leq T, \end{cases} \quad (2)$$

где $g(x,y)$ значение элемента выходного бинаризованного (двоичного) изображения в x -столбце y -строке, которое может

принимать только 2 значения (черное и белое без промежуточных серых тонов), обычно обозначаемых как 0 и 1.

3.1 Экспериментальное подтверждение работы алгоритма бинаризации

Предположим, имеется исходное изображение, которое необходимо бинаризовать (рис.8). Бинаризацию выполним пороговым методом.



Рисунок 8 – Исходное изображение

1. Определение размера изображения $M \times N$ пикселей и вычисление общего количества пикселей:

$$M = 329 \text{ пикселей}, \quad N = 615 \text{ пикселей},$$

$$M \times N = 329 \times 615 = 202335 \text{ пикселей}.$$

2. Определение уровня яркости каждого пикселя исходного изображения $f(x,y)$.
3. Построение гистограммы по полученным значениям яркостей изображения (рис.9).

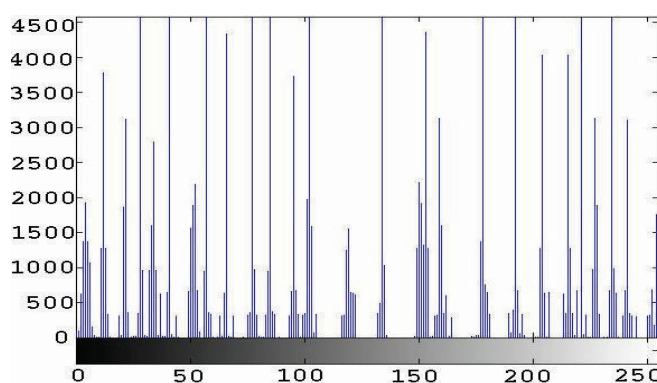


Рисунок 9 – Гистограмма уровней яркости пикселей изображения

4. На основе анализа гистограммы выбираем порог. Поскольку яркость пикселей исходного изображения распределены практически равномерно вдоль всей оси абсцисс гистограммы, то порог будет находиться посередине оси абсцисс, т.е.

$$T = 127,5.$$

Выполним сравнение значения яркости каждого пикселя с пороговым значением яркости в соответствии с системой уравнений (2) и получим бинаризованное изображение (рис.4).

Заключение

Научная новизна данной статьи заключается в том, что были показаны преимущества представления исходного изображения в градациях серого при выполнении пороговой бинаризации. Такой подход обладает рядом преимуществ перед изображениями в градациях базовых цветов RGB-модели. Была представлена возможность отображения семи шкал оттенков RGB-модели в виде шестиконечной звезды. Проведен анализ работы способа пороговой бинаризации исходного изображения, представленного в оттенках серого цвета. Работа предложенного метода даёт удовлетворительные результаты в том случае, если изображение состоит из пикселей с большой разницей уровней яркости, что позволяет получить бимодальное распределение на гистограмме, и тем самым, наиболее верно определить порог бинаризации. В случае размытых изображений с большим количеством оттенков серого, порог выбирается посередине всего интервала, на котором располагаются уровни яркости изображения.

Практическая значимость предложенного способа заключается в возможности выполнения бинаризации изображения в оттенках различных цветов (для шести базовых и серого) для решения задач обработки изображения в картографии, медицине, горнорудной промышленности, а также в спорте.

ЛИТЕРАТУРА

1. Прэтт У. Цифровая обработка изображений. - М.: Мир, 1982.- 309с.
2. Gonzales R., Wintz P. Digital image processing. Second edition. - Addison-wesley, 1987.
3. Trier and Torfinn Taxt. TECHNICAL REPORT: Evaluation of Binarization Methods for Document Images, Oslo, 1994