

Принцип получения оригинальной 3D-модели внешней поверхности кристалла в декартовой системе координат (рисунок 1) основывается на математическом описании объекта по определенному алгоритму с целью последующего построения поверхности кристалла [2].

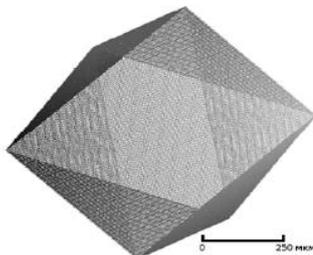


Рис. 1 – Пример 3D-модели кристалла алмаза в декартовой прямоугольной системе координат

Система (экспериментальная установка) содержит следующие составные части:

1. Держатель для необработанного объекта (кристалла сложной формы);
2. Систему сканирования, используемую для автоматического получения информации о геометрических параметрах и размерах объекта (кристалла сложной формы);
3. Систему освещения кристалла в виде двух (или более) лазерных пучков, направляемых по двум отдельным оптическим траекториям;
4. Систему формирования изображения кристалла сложной формы;
5. Вычислительное средство, реализующее возможность определения формы кристалла на основе принятого излучения от освещающей системы и указанной начальной информации о геометрической форме кристалла.

Устройство выполнено с возможностью приведения сканируемого объекта во вращение относительно осветительной системы [3] для предоставления пользователю более точной визуальной модели объекта.

Используя математическое описание, на основе знания пространственного расположения элементов системы освещения объекта, данные о показателе преломления исследуемых образцов кристаллов, можно построить трехмерные модели реальных кристаллов сложной геометрической формы. Полученный набор данных позволяет построить 3D виртуальную модель внешней поверхности кристалла и 3D виртуальные модели одного или нескольких видимых включений во внутреннем объеме кристалла. Эти данные могут быть использованы для последующего компьютерного анализа с целью предоставления пользователю информации, относящейся к визуальным характеристикам кристаллов сложной формы, которая может быть получена из исходного кристалла.

Разработка или усовершенствование существующих алгоритмов и систем определения геометрических параметров кристаллов сложной формы, являющихся целью представленной работы, существенно могут повысить эффективность алмазного производства, в том числе и ювелирного.

Список использованных источников:

1. Горбунова, Е.В. Возможности сортировки алмазного сырья опико-электронными методами / Е.В. Горбунова // Научно-технический вестник информационных технологий, механики и оптики. – СПб.: ИТМО, 2012. – №4 – С. 13 – 17.
2. Григорьев, Ю.М. Сравнительный анализ оптических свойств бриллиантов / Ю.М. Григорьев // Наука и образование. – М.: Наука, 2014. – №1 – С. 45 – 49.
3. Abraham Kerner., System for accurate 3d modeling of gemstones. PatentUS, no. 20160004926A1, 2012.

ДЕТЕКТИРОВАНИЕ ЛИЦ НА ИЗОБРАЖЕНИИ С ОДНОВРЕМЕННОЙ КЛАССИФИКАЦИЕЙ ЛИЦ С ПОМОЩЬЮ СВЁРТОЧНЫХ НЕЙРОННЫХ СЕТЕЙ

*Белорусский государственный университет информатики и радиоэлектроники
г. Минск, Республика Беларусь*

Суша А.В.

Вашкевич М.И. – к.т.н., доцент

Задача детектирования лиц является одной из самых важных в области компьютерного зрения, т.к. детектирование необходимо практически всем системам, которые оперируют изображениями человека. В работе рассматривается применение перспективного подхода детектирования лиц или объектов YOLO (You Only Look Once) на основе базовых моделей: VGG-16 и NasNet.

В настоящее время существует множество способов детектирования лиц. Большинство из них базируется на классических методах компьютерного зрения и машинного обучения (каскады Хаара,

HOG+SVM). Неоспоримым преимуществом этих подходов является высокая скорость работы. Однако данные методы имеют такие недостатки, как большое число ложных срабатываний и ограниченное число ракурсов лиц, на которых возможно проведение детектирования. Применение методов глубокого обучения позволяет решить эти недостатки. А более низкая скорость работы в настоящее время компенсируется наличием более производительных платформ.

Описываемый в данной работе метод детектирования лиц основывается на методе YOLO, описанный в [1] и усовершенствованный в [2]. Привлекательность метода YOLO заключается в том, что изображение анализируется свёрточной нейронной сетью только один раз, в отличие от подходов, базирующихся на применении скользящего окна. Также YOLO позволяет проводить некоторую дополнительную классификацию детектируемых объектов, например пол человека, выражение лица. В этих работах в качестве базовой модели использовалась архитектура сети *DarkNet*, разработанная авторами YOLO. Ключевым отличием настоящей работы является применение подхода YOLO к иным базовым моделям: VGG-16 [3] и *NasNet* [4] и своя архитектура *FaceNet*. Также некоторые отличия имеются в применяемой для обучения функции.

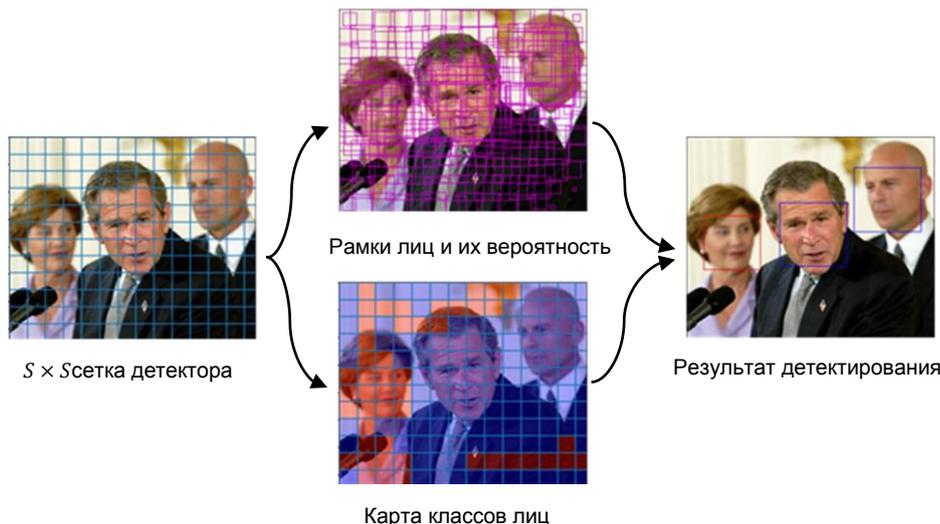


Рисунок 1 – Процесс детектирования объектов с помощью YOLO

На рисунке 1 изображён процесс детектирования объектов с помощью YOLO. Входное изображение разбивается на ячейки некоторого размера, которые формируют сетку размером $S \times S$. YOLO осуществляет предсказание как позиции рамки объекта, так и класса, к которому этот объект принадлежит. В общем виде процесс детектирования записывается следующим выражением:

$$B = transform(Y) = transform(D(I)),$$

где I – входное изображение;

D – детектор;

Y – выход детектора;

$transform$ – функция преобразования выхода сети в список объектов;

B – список найденных объектов:

$$B = [\{P_d, b_x, b_y, b_w, b_h, c\}],$$

где P_d – вероятность обнаружения объекта, $P_d = sigmoid(Y_p)$;

b_x, b_y – смещение по горизонтали и вертикали центра рамки объекта, в относительных координатах внутри ячейки, $b_x = sigmoid(Y_{b_x})$ и $b_y = sigmoid(Y_{b_y})$;

b_w, b_h – ширина и высота рамки в нормированных координатах относительно размера изображения, $b_w = \exp(Y_{b_w})$ и $b_h = \exp(Y_{b_h})$;

c – класс объекта, $c = argmax(softmax(Y_c))$.

В качестве функции ошибки для минимизации использовалась сумма функций ошибок для каждой величины, описывающей объект: для вероятности обнаружения объекта – средняя бинарная перекрёстная энтропия, для рамки лица – сумма средних квадратических ошибок для координат центра рамки и логарифмов ширины и высоты рамки и для класса объекта – средняя категориальная перекрёстная энтропия. Общая функция ошибки записывается следующим образом:

$$L_p = \frac{1}{N} \sum_{i=1}^N -d_i \cdot \log(\hat{d}_i) - (1 - d_i) \cdot \log(1 - \hat{d}_i) +$$

$$+ \frac{1}{|D|} \sum_{i \in D} (b_x - \hat{b}_x)^2 + (b_y - \hat{b}_y)^2 + (\log(b_w) - \log(\hat{b}_w))^2 + (\log(b_h) - \log(\hat{b}_h))^2 +$$

$$+ \frac{1}{|D|} \sum_{i \in D} \sum_{j=1}^M -c_{ij} \cdot \log(\hat{c}_{ij}),$$

где N – число примеров;

d_i – целевое значение вероятности обнаружения объекта i -го примера;

\hat{d}_i – предсказанное значение вероятности обнаружения объекта i -го примера;

D – множество индексов примеров, для которых при $d_i = 1$, $\hat{d}_i \geq 0,5$;

b_x, b_y, b_w, b_h – целевые значения смещения центра и размера рамки объекта i -го примера;

$\hat{b}_x, \hat{b}_y, \hat{b}_w, \hat{b}_h$ – предсказанные смещения центра и размера рамки объекта i -го примера;

M – число классов объектов;

t_{ij} – целевое значение вероятности принадлежности объекта к j -го классу i -го примера;

y_{ij} – предсказанное значение вероятности принадлежности объекта к j -го классу i -го примера.

Обучение проводилось на базах данных *FaceSrub* [5] и *IMDB* [6]. Несмотря на наличие более 350 тысяч обучающих примеров, в указанных базах размечена рамка лица только одного человека даже если людей изображено несколько. Но как оказалось даже в таких условиях свёрточные сети способны обучаться. Для оценки качества обучения использовалась база *Fddb* [7] содержащая 2845 вручную размеченных изображений лиц. Результат тестирования приведен в таблице 1. В качестве метрик использовались точность, полнота, *f1-score* и также пересечение над объединением (англ. *IoU* – *intersection over union*). Пороговым значением вероятности детектирования – 0,5; Пороговое значение *IoU* для алгоритма подавления не максимумов – 0,4 и для фиксирования правильно сдетектированного лица – 0,5.

Таблица 1 – результаты тестирования

Базовая модель	Точность, %	Полнота, %	F1-Score, %	IoU, %
VGG-16	12,02	11,01	11,49	60,59
NasNet	99,07	85,84	91,98	78,07
FaceNet	88,70	80,72	84,52	70,28

Исходя из приведённых результатов можно заключить, что VGG-16 обучилась довольно плохо. К тому же время, затраченное на обучение этой сети, в несколько раз больше времени, затраченного на обучение *NasNet*. С одной стороны *NasNet* имеет меньшее число параметров, но с другой – более сложную архитектуру. Это хорошо показывает, что архитектура нейронной сети оказывает существенное влияние на качество её работы. В настоящее время проводится работа по созданию и обучению собственной архитектуры базовой модели – *FaceNet*. Основной акцент при проектировании делается на возможности применения этой модели на мобильных вычислительных платформах с сохранением высокой точности детектирования.

Список использованных источников:

1. Joseph Redmon, Santosh Divvala, Ross Girshick, Ali Farhadi, You Only Look Once: Unified, Real-Time Object Detection // arXiv preprint, arXiv:1506.02640. – 2015. – 10 p.
2. Joseph Redmon, Ali Farhadi, YOLO9000: Better, Faster, Stronger // arXiv preprint, arXiv:1612.08242 – 2016. – 9 p.
3. Karen Simonyan, Andrew Zisserman, Very Deep Convolutional Networks for Large-Scale Image Recognition // arXiv preprint, arXiv:1409.1556. – 2014. – 14 p.
4. Barret Zoph, Vijay Vasudevan, Jonathon Shlens, Quoc V. Le, Learning Transferable Architectures for Scalable Image Recognition // arXiv preprint, arXiv:1707.07012. – 2017. – 14 p.
5. Hong-Wei Ng, Stefan Winkler, A Data-driven approach to cleaning large face datasets // In Proc. IEEE International Conference on Image Processing – 2014. – pp. 27-30.
6. Rasmus Rothe, Radu Timofte, Luc Van Gool, Deep expectation of real and apparent age from a single image without facial landmarks // International Journal of Computer Vision. – 2016. – pp. 144-157.
7. Vidit Jain, Erik Learned-Miller, Fddb: A Benchmark for Face Detection in Unconstrained Settings // Technical Report UM-CS-2010-009, Dept. of Computer Science, University of Massachusetts. – 2010. – 11 p.

СИСТЕМЫ ФАЗОВОЙ АВТОПОДСТРОЙКИ ЧАСТОТЫ

Белорусский государственный университет информатики и радиоэлектроники
г. Минск, Республика Беларусь

Дорошкевич М.В., Зайцева А.И., Фоменок А.А.

Качинский М. В. – к.т.н., доцент

Вопросы синхронизации требуют особого внимания, когда речь идет о распространении цифровой информации. При использовании фазоманипулированных сигналов (phase-shiftkeying (PSK)) необходимо различать передаваемые символы по фазе. Однако опорные генераторы на передающей и приемной стороне не могут быть согласованы. Для решения этой проблемы используются следящие контуры (петли), основная задача которых производить подстройку генераторов опорных сигналов для когерентной демодуляции. Такие системы получили название ФАПЧ.

Система фазовой автоподстройки частоты — система автоматического регулирования, которая подстраивает фазу управляемого генератора так, чтобы достиглось равенство с фазой опорного сигнала,