

Нейросетевая модель распознавания поверхностей машиностроительных деталей по результатам оптического сканирования

М.А. Болотов¹, В.А. Печенин¹, Н.В. Рузанов¹

¹Самарский национальный исследовательский университет им. академика С.П. Королева, Московское шоссе 34А, Самара, Россия, 443086

Аннотация. Для прогнозирования показателей качества изделий (в частности, сборочных параметров) используются математические модели, реализованные в виде компьютерных моделей. Для адекватности расчётов необходимо иметь информацию о действительной геометрии деталей, которую можно получить с помощью бесконтактных измерений деталей сборки. В результате измерения деталей и узлов при помощи оптических или лазерных сканеров формируется массив измеренных точек большой размерности. После проведения стандартной обработки данных (например удаление шума, совмещение сканов, сглаживание, создание триангуляционной сетки) возникает необходимость распознавания отдельных поверхностей деталей. В работе приведена нейросетевая модель, позволяющая выполнять распознавание элементов на основе использования массива измеренных точек, полученных посредством сканирования.

1. Введение

Наименее автоматизированным этапом в промышленности является сборка единичных и серийных изделий, характеризующихся средней и высокой сложностью. К таким изделиям относятся авиационные двигатели. Рассматриваемые изделия не производятся массово как автомобили, характеризуются высокой степенью опциональности и повышенными требованиями к сложности и точности. Доля трудоёмкости сборки в общей трудоёмкости изделий составляет до 25% и в значительной мере определяет их качество. Существует несколько причин, затрудняющих полную автоматизацию сборки названных изделий. Одна из существенных причин заключается в сложности определения параметров выполнения операции роботами, гарантированно обеспечивающих заданные точность и качество изделий. Сборка изделий средней и высокой сложности является уникальной операцией, в ходе которой по результатам измерений и анализа геометрии собираемых деталей изменяется ход операций. Измерение геометрии производится бесконтактными и контактными способами.

Для частичной автоматизации процессов сборки двигателя необходимо распознавать как отдельные детали, так и поверхности деталей, по которым будет проходить сборка. Распознавание граней возможно за счёт использования подходов компьютерного зрения [1,2,3]. Целью работы является создание модели, основанной на использовании нейронных сетей, предназначенной для распознавания поверхностей машиностроительных деталей после их измерения на оптическом или лазерном сканерах.

обучения поверх таких описаний трёхмерных фигур создаёт ряд проблем. Во-первых, размер организованных баз данных с аннотированными 3D-моделями довольно ограничен по сравнению с наборами данных изображений. Например, ModelNet содержит около 150 тыс. объектов. Напротив, база данных ImageNet [14] уже включает десятки миллионов аннотированных изображений. Во-вторых, дополнительные признаки и метрики трехмерных фигур имеют тенденцию быть очень многомерными, что делает алгоритмы склонными к переобучению.

Одной из последних работ для задач классификации объектов и сегментации их отдельных частей, в которой на входе в сеть используется непосредственно массив измеренных точек, является работа [15]. Разработанная авторами из Стэнфордского университета сеть названа PointNet, основная идея подхода – запомнить пространственные признаки каждой точки и затем объединить все индивидуальные признаки точек в промаркированное общее облако точек. В основе сети лежит свёрточная архитектура. Основным минусом приведенной сети является то, что на входе необходимо иметь всегда одинаковое количество точек для всех объектов, что в практике не возможно и придется прибегать к искусственному «проряжению» данных.

На основе проведённого обзора литературы и специфики решаемой задачи было принято решение в текущей задаче использовать при решении подход на основе использования 2D-проекций объектов. В данном подходе используется свёрточная нейронная сеть для сегментации U-net [16].

Раскроем основную идею и этапы разработанной модели для сегментации отдельных граней. Основная идея заключается в создании проекций (картинок) для фасетной модели измеренных данных, сегментация граней на проекциях, идентификация фасет по сегментируемым изображениям. Этапы подхода приведены на рисунке 1.

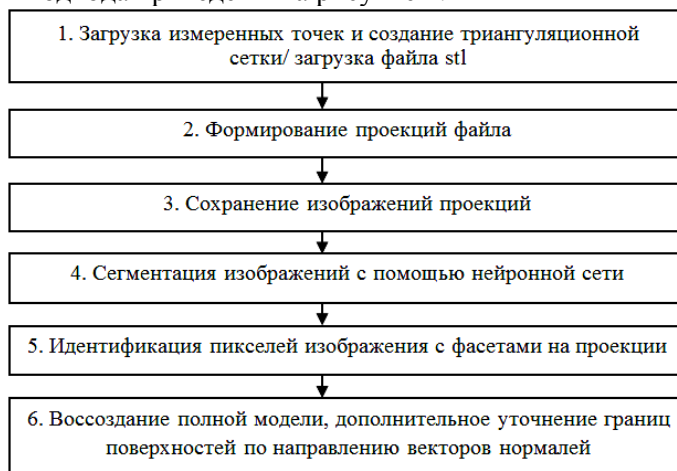


Рисунок 1. Блок-схема модели распознавания граней с помощью проекций измеренных данных.

Рассмотрим этапы распознавания граней более подробно.

3.1. Загрузка измеренных точек и создание триангуляционной сетки/ загрузка файла stl

Как отмечалось во введении, модель предназначена для распознавания геометрии после измерения на оптических и лазерных сканерах. После измерения на сканере и проведения предварительной обработки данных создаётся файл с координатами точек, объединённых в фасетную поверхность формата *.stl. В файле содержатся следующие данные: $V_{g \times 3}$ (матрица координат вершин сетки stl-модели), $F_{m \times 3}$ (матрица сочетаний вершин по три, образующих фасеты поверхностей), $N_{m \times 3}$ (матрица координат нормалей фасет).

3.2. Формирование проекций файла

Для возможности семантической сегментации фасет на отдельные поверхности с помощью глубоких нейронных сетей необходимо создание проекций 3D-поверхностей на координатные плоскости. Для подготовки проекций был использован алгоритм Робертса [17].

Алгоритм Робертса представляет собой первое известное решение задачи об удалении невидимых линий. Это математически элегантный метод, работающий в объектном пространстве. Алгоритм прежде всего удаляет из каждого тела те ребра или грани, которые экранируются самим телом. Затем каждое из видимых ребер каждого тела сравнивается с каждым из оставшихся тел для определения того, какая его часть или части, если таковые есть, экранируются этими телами. Поэтому вычислительная трудоемкость алгоритма Робертса растет теоретически, как квадрат числа объектов.

Работа алгоритма Робертса проходит в два этапа:

1. Определение нелицевых граней для каждого тела отдельно.
2. Определение и удаление невидимых ребер.

Для подготовки данных был использован только первый этап алгоритма. Второй этап не столь необходим для дальнейшего решения, имеет большую сложность и требует дополнительно созданных фасет.

При создании проекций на координатные оси получаются ортогональные проекции в визуальной плоскости. Например, при проекции на плоскость XOY , сохраняется структура $F_{m \times 3}$, а матрица вершин $V_{g \times 3}$ преобразуется в $V_{g \times 2}$, имеющей лишь координаты по осям x и y . Обходя вершины полученных проекций фасетов в той же последовательности, что и в оригинале, можно разделить их на два вида: те, что ориентированы против хода часовой стрелки, что означает что мы смотрим на фасету с внешней стороны тела; те, что ориентированы по ходу часовой стрелки, что означает что мы смотрим на фасету изнутри. Порядок следования вершин определяет направление нормали. Таким образом, если компонента вектора нормали плоскости проекции (в рассматриваемом примере компонента по оси z) имеет отрицательную величину, мы смотрим на фасету изнутри. Так как объект ограничен замкнутой поверхностью, то мы не можем наблюдать грани с внутренней стороны – они являются невидимыми. Таким образом, из структуры $F_{m \times 3}$ необходимо исключить выявленные по приведенному выше методу фасеты, получив структуру проекции $F_{m_1 \times 3}$, где $m_1 < m$.

3.3. Сохранение изображений проекций

Сохранение изображения производится с помощью библиотек работы с stl (функции пакета stltools автора Pau Micó) и средств сохранения графики MATLAB. Размеры картинок сохраняются и одинаковы для всех проекций. С помощью этих же инструментов сохраняются обучающие картинки проекций, где объекты для распознавания выделены различными цветами (рисунки 2 и 3).

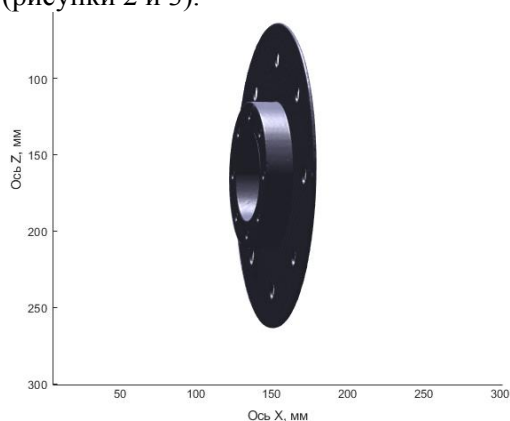


Рисунок 2. Stl-файл детали «имитатор проставки».

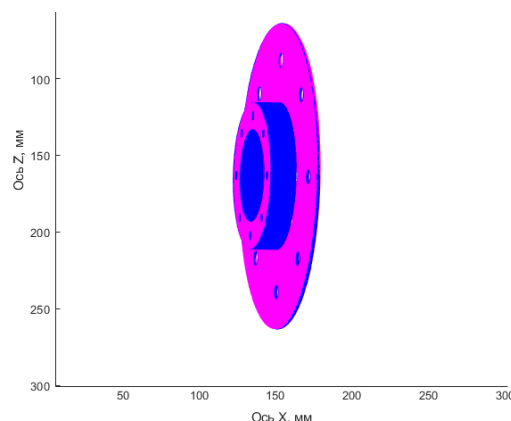


Рисунок 3. Проекция детали, используемая для сегментации.

Тем не менее для сегментации требуются четкие границы, поэтому сохраненные изображения в формате *.png импортируются в MATLAB (в переменную загружаются матрицы RGB-цветов), изображение переводится в оттенки серого с использованием выражения:

$$I_{сер} = 0,299 \cdot R + 0,587 \cdot G + 0,114 \cdot B, \quad (1)$$

где $I_{сер}$ – матрица изображения «в оттенках серого»; $\mathbf{R}, \mathbf{G}, \mathbf{B}$ – матрицы-компоненты системы RGB.

Сохраняются многомерные матрицы изображений в файл *.mat. Всего сохраняется шесть проекций. Объект помещён как бы в куб, грани которого параллельны плоскостям координат. Размеры куба таковы, что он включает в себя все измеренные объекты выборки. Проекции соответственно делаются на грани куба: две на грани параллельные плоскости XOY ; две, параллельные плоскости XOZ и две параллельные YOZ .

3.4. Сегментация изображений с помощью нейронной сети

Архитектура сети U-Net приведена на рисунке 4.

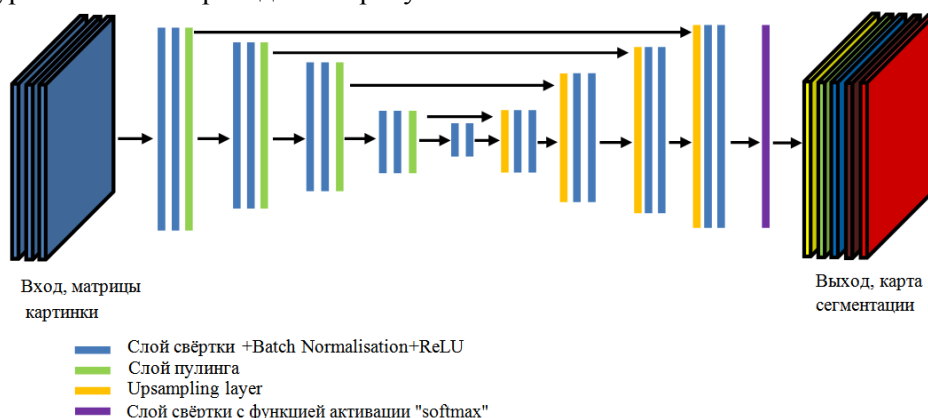


Рисунок 4. Архитектура нейронной сети U-Net для сегментации изображения.

Архитектура сети представляет собой последовательность слоёв свёртка+пулинг, которые сначала уменьшают пространственное разрешение картинки, а потом увеличивают его, предварительно объединив с данными картинки и пропустив через другие слои свёртки. Таким образом, сеть выполняет роль своеобразного фильтра.

Первая половина сети содержит слои свертки с функцией активации ReLU, нормализацию по мини-батчам и слоями пулинга (субдискретизации) и её можно назвать сжимающим путём. Вторая часть представляет собой расширяющий путь.

Слой Upsampling – слой, обратный пулингу, который расширяет карту признаков, после которого следует свёртка, уменьшающая количество каналов признаков. После идёт конкатенация («склеивание» линейных объектов) с соответствующим образом обрезанной картой признаков из сжимающего пути и два свёрточных слоя.

На последнем слое свёртка с ядром 1×1 используется для приведения каждого 64-компонентного вектора признаков до требуемого количества классов. Функция активации на последнем слое – «softmax».

Сеть была воспроизведена в программной среде Python.

3.5. Идентификация пикселей изображения с фасетами на проекции

После проведения сегментации изображения производится сопоставление его пикселей координатам соответствующей проекции. Для сопоставления пиксели переводятся в точки на соответствующей грани куба, описанного в разделе 3.3. Значения пикселей изображения лежат в интервале $[0; 255]$. Соответственно для идентификации пикселей, отвечающих за определенную грань, в матрице изображения ищутся значения определенной интенсивности. При этом поиск производится с определённым допуском. Положение найденных пикселей в матрице (строка-столбец) переводятся с учётом масштаба в координаты пространства (две координаты с изображения, третья - координата соответствующей грани куба). При этом в разделе 3.2 не были удалены грани, закрываемые другими. Поэтому при сопоставлении точка объекта на изображении может попасть на две и более фасеты проекции. В этом случае выбирается фасета, ближайшая к точке вида.

3.6. Воссоздание полной модели, дополнительное уточнение границ поверхностей по направлению векторов нормалей

Предыдущие шаги были необходимы для предварительного автоматизированного распознавания лишь части измеренных точек. После идентификации вершин и фасетов разным поверхностям производится сопоставление их с полным набором вершин (поиск по равенству координат точек). Выбираются точки, приближенные к геометрическому центру поверхности. По вершине в геометрическом центре выбирается фасета, принадлежащая определенной грани тела. На следующем этапе используется алгоритм поиска и уточнения фасет одной грани. Кратко опишем этапы алгоритма поиска фасетов, принадлежащих определенной грани [18]:

1) на входе задаются параметры $\mathbf{V}_{g \times 3}$, $\mathbf{F}_{m \times 3}$, $\mathbf{N}_{m \times 3}$, \vec{p} (вектор координат точки на поверхности, фасеты которой необходимо выделить), \vec{t} (нормаль в точке p), α (допуск на угол для поиска соответствующих фасет).

2) производится поиск фасеты, которой принадлежит точка p (точка пересечения луча и фасеты).

3) производится поиск всех фасет, у которых вектор нормали отличается от вектора нормали \vec{t} не более чем на величину α (для проверки углов используется скалярное произведение нормальных векторов).

Из всех найденных фасет, подходящих по направлению нормали, выбираются связанные с первой и между собой общими вершинами.

4. Результаты распознавания

Для измерения детали использовался оптический 3D сканер RANGEVISION Pro2M. На рисунках 5 и 6 представлены фотографии процесса измерения детали «имитатор проставки» (чертёж на рисунке 1).



Рисунок 5. Измерение одного скана детали.



Рисунок 6. Измерение группы сканов детали.

Для оценки качества сегментации можно использовать видоизмененную функцию потерь, приведенную в работе [19]. В работе предлагается вычислять ошибку определения параметров прямоугольника при распознавании лиц и изображений на картинке с помощью метрики Intersection over Union (IoU), равной отношению площади пересечения прямоугольника, полученного в результате детектирования, и прямоугольника из разметки к площади их объединения.

В нашем случае происходит работа с фасетами тела, поэтому вместо площадей можно оперировать количеством фасет. Поэтому рассчитывается коэффициент $\delta_{сегм}$, равный отношению количества фасет $N_{p \cap \partial}$, являющихся пересечением множества фасет поверхности, полученного в результате распознавания, и действительно принадлежащей данной поверхности фасет к множеству фасет, являющегося их объединением $N_{p \cup \partial}$:

$$\delta_{сегм} = N_{p \cap \partial} / N_{p \cup \partial}. \quad (2)$$

Таким образом, величина коэффициента лежит в интервале $[0; 1]$. Если распознаваемых поверхностей несколько, то можно рассчитывать обобщенный коэффициент $\delta_{сегм}$ для всех распознаваемых фасет тела.

Количество измеренных фасет детали m составило 622130. Для обучения нейронной сети была сформирована выборка из 1000 случаев stl-модели рассматриваемой детали, по-разному расположенных в пространстве, соответственно всего 6000 проекций. Для формирования обучающей выборки использовалась номинальная модель детали, сохранённая в формате *.stl. Распознанные плоские и цилиндрические грани измеренной на сканере детали «имитатор проставки» с использованием всех шести проекций приведены на рисунках 7 и 8. В таблице 1 приведены рассчитанные коэффициенты $\delta_{сегм}$ соответственно для цилиндрических, плоских и всех распознаваемых граней измеренной детали.

Таблица 1. О погрешности распознавания фасет граней детали.

Объект оценки	Цилиндрические грани	Плоские грани	Все распознаваемые грани
Значение $\delta_{сегм}$	0,76	0,95	0,91

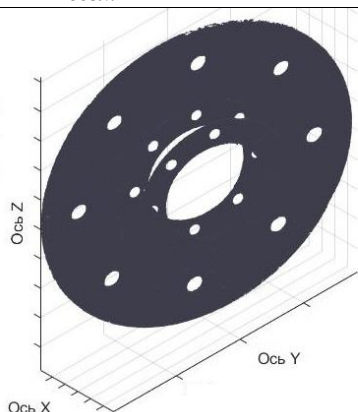


Рисунок 7. Распознанные плоские грани.

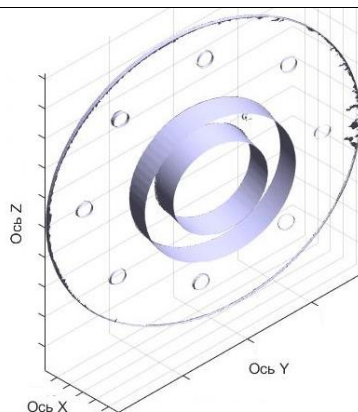


Рисунок 8. Распознанные цилиндрические грани.

Погрешность распознавания цилиндрических граней выше, чем плоских. Хотя все цилиндрические грани были идентифицированы, присутствовали фасеты, принадлежащие плоским граням, а так же немного фасет от резьбовых отверстий. Общий же коэффициент превысил 90%, что связано с гораздо большим количеством фасет плоских граней по сравнению с цилиндрическими.

5. Заключение

Приведённая в работе модель позволяет быстро производить распознавание цилиндрических и плоских поверхностей деталей, за счёт использования обученной нейронной сети и алгоритма поиска фасет грани по ранее распознанным фасетам. Разработанная модель нужна для дальнейшего прогнозирования сборочных параметров изделия, основанного на компьютерном моделировании [20]. Использование цифрового прогнозирования позволит роботам-сборщикам функционировать без участия человека.

6. Литература

- [1] Zakani, F.R. Segmentation of 3D meshes combining the artificial neural network classifier and the spectral clustering / F.R. Zakani, M. Bouksim, K. Arhid, M. Aboulfatah, T. Gadi // Computer Optics. – 2018. – Vol. 42(2). – P. 312-319. DOI: 10.18287/2412-6179-2018-42-2-312-319.
- [2] Блохинов, Ю.Б. Разработка алгоритма семантической сегментации аэрофотоснимков реального времени / Ю.Б. Блохинов, В.А. Горбачев, Ю.О. Ракутин // Компьютерная оптика. – 2018. – Т. 42, № 1. – С. 141-148. DOI: 10.18287/2412-6179-2018-42-1-141-148.
- [3] Никитин, М.Ю. Нейросетевая модель распознавания человека по лицу в видеопоследовательности с оценкой полезности кадров / М.Ю. Никитин, В.С. Конушин,

- А.С. Конушин // Компьютерная оптика. – 2017. – Т.41, №5. – С. 732-742. DOI: 10.18287/2412-6179-2017-41-5-732-742.
- [4] Николенко, С.И. Глубокое обучение / С.И. Николенко, А. Кадури, Е. Архангельская. – СПб.: Питер, 2018. – 480 с.
- [5] Hubei, D.H. Eye, Brain, and Vision / D. H. Hubei. – New York: Scientific American, 1988.
- [6] Webvision: The Organization of the Retina and Visual System [Electronic resource]. – Access mode: <http://webvision.med.utah.edu/book/>.
- [7] Wu, Z. 3D ShapeNets: A deep representation for volumetric shape modeling / Z. Wu, S. Song, A. Khosla, F. Yu, L. Zhang, X. Tang, J. Xiao // Proc. CVPR. – 2015. – P. 1-9.
- [8] Horn, B.K.P. Extended gaussian images / B.K.P. Horn // Proc. of the IEEE. – 1984. – Vol. 72(12). – P. 1671-1686.
- [9] Osada, R. Shape distributions / R. Osada, T. Funkhouser, B. Chazelle, D. Dobkin // ACM Transactions on Graphics. – 2002. – Vol. 21(4). – P. 807-832.
- [10] Kazhdan, M. Rotation invariant spherical harmonic representation of 3D shape descriptors / M. Kazhdan, T. Funkhouser, S. Rusinkiewicz // Proceedings of the 2003 Eurographics/ACM SIGGRAPH symposium on Geometry processing. – 2003. – P. 156-164.
- [11] Chaudhuri, S. Data-driven suggestions for creativity support in 3D modeling / S. Chaudhuri, V. Koltun // ACM Transactions on Graphics. – 2010. – Vol. 29(6). – P. 183. DOI: 10.1145/1866158.1866205.
- [12] Bronstein, A. Shape Google: Geometric words and expressions for invariant shape retrieval / A. Bronstein, M. Bronstein, M. Ovsjanikov, L. Guibas // ACM Transactions on Graphics. – 2011. – Vol. 30(1). – P. 1. DOI: 10.1145/1899404.1899405.
- [13] Kokkinos, I. Intrinsic shape context descriptors for deformable shapes / I. Kokkinos, M. Bronstein, R. Litman, A. Bronstein // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2012. DOI: 10.1109/CVPR.2012.6247671.
- [14] Deng, J. Imagenet: A large-scale hierarchical image database / J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, L. Fei-Fei // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2009. DOI: 10.1109/CVPR.2009.5206848.
- [15] Qi, C.R. PointNet: Deep learning on point sets for 3d classification and segmentation / C.R. Qi, H. Su, K. Mo, L.J. Guibas // arXiv preprint arXiv:1612.00593. – 2016. – P. 652-660.
- [16] Ronneberger, O. U-net: Convolutional networks for biomedical image segmentation / O. Ronneberger, P. Fischer, T. Brox // Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). – 2015. – Vol. 9351. – P. 234-241.
- [17] Дёмин, А.Ю. Компьютерная графика / А.Ю. Дёмин, А.В. Кудинов. – Томский политехнический университет, 2005 [Электронный ресурс]. – Режим доступа: <http://compgraph.tpu.ru/roberts.htm> (01.12.2018).
- [18] Степаненко, И.С. Методика повышения точности деталей ГТД, изготавливаемых селективным лазерным спеканием / И.С. Степаненко, В.А. Печенин, Н.В. Рузанов, А.И. Хаймович // Сборник трудов ИТНТ-2018. – Самара: Новая техника, 2018. – С. 1672-1680.
- [19] Yu, J. UnitBox: An Advanced Object Detection Network / J. Yu, Y. Jiang, Z. Wang, Z. Cao, T. Huang // INNS, 2016. DOI: 10.1145/2964284.2967274.
- [20] Болотов, М.А. Метод оценки неопределённостей пространственного сопряжения высокоточных оптических и механических деталей / М.А. Болотов, В.А. Печенин, С.П. Мурзин // Компьютерная оптика. – 2016. – Т. 40, № 3. – С. 360-369.

Благодарности

Работа поддержана Министерством образования и науки Российской Федерации в рамках реализации программы повышения конкурентоспособности Самарского университета среди мировых ведущих научно-образовательных центров на 2013-2020 годы. Экспериментальные исследования были выполнены на оборудовании ЦКП САМ-технологий (RFMEFI59314X0003).

Neural recognition model surfaces of machine parts based on the results of the optical scanning

М.А. Bolotov¹, В.А. Pechenin¹, N.V. Ruzanov¹

¹Samara National Research University, Moskovskoe Shosse 34A, Samara, Russia, 443086

Abstract. To predict the quality parameters of products (in particular, the assembly parameters) used mathematical models implemented in the form of computer models. For the adequacy of calculations, it is necessary to have information about the actual geometry of the parts, which can be obtained using noncontact measurements of parts of the assembly. As a result of measurement of parts and components using optical or laser scanner is formed an array of measured points large dimension. After the standard processing (e.g. noise removal, combining the scans, smoothing, creating triangulation mesh) becomes necessary recognition of individual surfaces of parts. The paper presents a neural network model that allows the recognition of elements based on the use of an array of measured points obtained by scanning.