

НЕЙРОКАЛИБРОВКА СТЕРЕОПАРЫ

Abstract: The task of stereopair calibration using neural networks is considered. Different methods of neurocalibration are investigated which differ in number of hidden layers, number of neural modules, input data preprocessing, activation function and learning algorithm. Unified procedure for finding optimal number of hidden neurons is developed. Optimal number of neurons in hidden layers is determined for each of the neurocalibration methods according to this procedure. Neurocalibraion methods are compared with analytical ones on synthetic and real data. Practical recommendations for neurocalibration of stereopairs with low distorsion are formulated.

Key words: calibration, stereopair, neural networks.

Анотація: Розглянуто задачу калібровки стереопари за допомогою нейронних мереж. Досліджено різні методи нейрокалібрування, які відрізняються кількістю прихованих шарів, кількістю нейромережових модулів, препроцесінгом вхідних даних, функцією активації та алгоритмом навчання. Розроблено уніфіковану методику знаходження оптимальної кількості нейронів у прихованих шарах та визначено оптимальну кількість нейронів у прихованих шарах для кожного з методів. Проведено порівняння методів нейрокалібровки між собою та з аналітичними методами калібровки на синтетичних та реальних даних. Сформульовано практичні рекомендації щодо процедури нейрокалібровки стереопар зі слабкими нелінійними спотвореннями.

Ключові слова: калібрування, стереопара, нейронні мережі.

Аннотация: Рассмотрена задача калибровки стереопары при помощи нейронных сетей. Исследованы различные методы нейрокалибровки, отличающиеся количеством скрытых слоев, количеством нейросетевых модулей, препроцессингом входных данных, функцией активации и алгоритмом обучения. Разработана унифицированная методика нахождения оптимального количества нейронов в скрытых слоях и определено оптимальное количество нейронов в скрытых слоях для каждого из методов. Проведено сравнение методов нейрокалибровки между собой и с аналитическими методами калибровки на синтетических и реальных данных. Сформулированы практические рекомендации по процедуре нейрокалибровки стереопар со слабыми нелинейными искажениями.

Ключевые слова: калибровка, стереопара, нейронные сети.

1. Введение

Для решения большинства задач стереозрения требуется предварительная калибровка стереопары. Классический подход к калибровке стереопары при известных 3D-координатах калибровочных точек состоит в отдельной калибровке каждой из камер и опирается на хорошо изученные модели и методы калибровки одной камеры [1–4]. Мы предлагаем новый подход с использованием нейронных сетей. Такая постановка задачи выглядит вполне естественно на фоне общего развития нейроматематики [5] и простоты переформулирования задачи калибровки как задачи обучения или задачи аппроксимации функции. К преимуществам нейронных сетей перед аналитическими методами можно отнести:

– нейронные сети в состоянии аппроксимировать гораздо более широкий класс функций, чем общепринятая в современной калибровке линейная модель камеры с поправками на дисторсию в виде степенного ряда [2–4]. Нейронные сети могут учитывать практически любые нелинейные искажения;

– при нейрокалібровке не требуется знание математической модели камеры. Напротив, при аналитической калибровке неправильный выбор модели камеры может сильно повлиять на результаты калибровки. Слишком простая модель дает большую погрешность, слишком сложная – приводит к проблемам многомерной оптимизации и сложна в реализации;

– нейросетевая калибровка не зависит от начального приближения. Учет дисторсионных искажений в аналитических моделях камеры приводит к многомерной оптимизации, которая в

большинстве случаев чувствительна к точности начального приближения. Методы нахождения начального приближения составляют отдельную область аналитической калибровки и не всегда дают удовлетворительный результат. Как правило, для нахождения начального приближения используется линейная модель камеры, а это заведомо неверно в случае камер с существенной дисторсией;

- нейронные сети обучаются непосредственно на 3D ошибку триангуляции, в то время как аналитические методы калибруют каждую камеру отдельно на 2D ошибку проектирования и затем используют найденные параметры для триангуляции. Причиной отдельной калибровки камер опять же служат трудности многомерной оптимизации. В редких случаях параметры камер затем уточняются по ошибке триангуляции;

- нейронные сети выступают в роли «черного ящика». Экспериментатору необходимо только подготовить обучающую выборку, от него не требуются знание модели камеры, численных методов, а также разработка и реализация специализированных программ для калибровки. В силу дешевизны современных цифровых камер задачами стереокалибровки начинает заниматься широкий круг неспециалистов в области калибровки, и для них подход «черного ящика» может быть существенным плюсом;

- нейрокалибровка может быть удобна в системах распознавания образов и машинного зрения, построенных на основе нейронных сетей, так как позволяет унифицировать архитектуру системы;

- быстродействие нейронных сетей в рабочем режиме потенциально выше, чем у любого аналитического метода, за счет естественного распараллеливания операций в нейронных сетях;

- нейронные сети устойчивы к шумам и погрешностям в данных и не требуют знания их вероятностных характеристик.

Нейросетевой подход к задаче нейрокалибровки стереопары был предпринят в [0]. Однако автор остановился на простейшей архитектуре сети, аргументируя это ресурсоемкостью обучения. Соответственно при обучении нейронных сетей непосредственно на ошибку триангуляции была получена большая ошибка по сравнению с другими методами, и дальше этот вопрос не рассматривался. Вместо этого была построена схема обучения нейронных сетей на ошибку линейного метода, работу которой нам воспроизвести не удалось. Кроме того, в работе использовались очень небольшие обучающие выборки, по количеству векторов сравнимые с количеством весовых коэффициентов обучавшихся нейронных сетей, что обычно приводит к локализации обучения [7, 8]. Также не была проведена проверка на реальных массивах данных.

В нашей работе заново рассматривается задача обучения нейронных сетей непосредственно на ошибку триангуляции при известных пространственных координатах калибровочных точек. Исследуются различные нейросетевые методы калибровки, отличающиеся препроцессингом, количеством скрытых слоев нейронной сети, количеством нейросетевых модулей, алгоритмом обучения и функцией активации. Рассматривается задача обучения нейронных сетей на ошибку линейного метода, описанная в [6]. Нейросетевые методы проверяются на реальных и синтетических данных.

В разделе 2 дается постановка задачи и описываются рассматриваемые методы нейрокалибровки. Раздел 3 посвящен общей для всех рассматриваемых подходов методике калибровки. Сравнение различных методов калибровки на реальных данных дается в разделе 4. Проверка полученных результатов на синтетических данных приводится в разделе 5. Раздел 6 содержит заключение и выводы, а также набор практических рекомендаций по нейрокалибровке стереопары.

2. Постановка задачи и рассматриваемые методы нейрокалибровки

Цели нашей работы:

- исследование применимости нейронных сетей к задаче калибровки стереопары в целом, сравнение разных методов нейрокалибровки между собой и с аналитическими методами по точности калибровки;
- поиск лучшего метода нейрокалибровки и выработка практических рекомендаций по нейрокалибровке;
- определение оптимальной архитектуры нейронных сетей для случая камер со слабыми нелинейными искажениями.

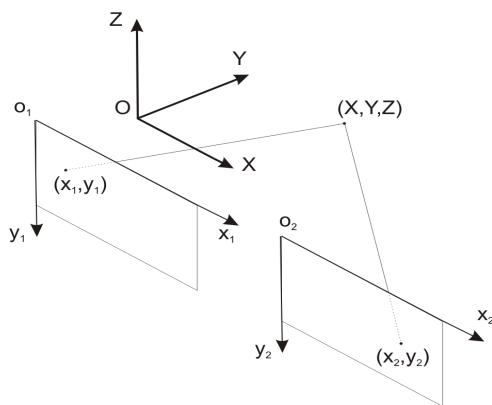


Рис. 1. Проектирование 3D точки на плоскости изображений двух камер стереопары

Калибровка стереопары заключается в нахождении параметров модели стереопары, которые позволяют восстанавливать 3D-координаты некоторой точки пространства по 2D-координатам ее изображений на двух снимках стереопары. В дальнейшем будем обозначать через (X, Y, Z) координаты точки в пространстве, через (x_1, y_1) и (x_2, y_2) – координаты ее изображения на первом и втором снимках стереопары. При этом в дальнейшем изложении предполагается, что ось O_X примерно параллельна сенсорам камер и отрезку, соединяющему их оптические центры, ось O_Y обозначает глубину и направлена перпендикулярно плоскости сенсоров, а ось O_Z определяет высоту точки над поверхностью (рис. 1).

Методы нейрокалибровки, которые рассматриваются в работе, можно классифицировать по следующим признакам:

1. Препроцессинг входного вектора:

- отсутствует. На вход нейронной сети подаются четыре действительных числа: x_1, y_1, x_2, y_2 (рис. 2а);
- на вход нейронной сети подаются три действительных числа. Два из них, как и в предыдущем случае, x_1 и y_1 . Третье число представляет собой разность x_1 и x_2 . Такой препроцессинг имеет смысл в силу эпилептического ограничения на координаты изображений точки

на снимках для линейных камер. Из эпиполярной геометрии известно, что для триангуляции достаточно трех из четырех компонент 2D-координат. Отбрасывание лишней компоненты снижает размерность входного вектора и сложность нейронной сети. Использование $(x_1 - x_2)$ в качестве третьего входа объясняется тем, что камеры были практически соосны, и именно эта величина важна для определения глубины (рис. 2б).

2. Количество модулей в сети:

- один с тремя действительными выходами: X , Y и Z (рис. 2а, 2б);
- три модуля, каждый с одним действительным выходом. При этом обучаются три нейронные сети, каждая на свою пространственную координату (рис. 2в). Снижение размерности выхода упрощает архитектуру нейронных сетей.

3. Количество скрытых слоев в нейронных сетях [7, 8]:

- один. Нейронные сети с одним скрытым слоем являются универсальным аппроксиматором достаточно широкого класса функций, поэтому исследование начинается именно с таких сетей;
- два. Нейронные сети с двумя скрытыми слоями могут аппроксимировать функции с разрывами, кроме того, они могут быть гораздо эффективнее сетей с одним скрытым слоем за счет уменьшения необходимого числа нейронов в каждом из слоев;
- три. Дальнейшее увеличение количества скрытых слоев позволяет уменьшить количество нейронов в каждом слое и задает более сложную разделяющую поверхность. Увеличение числа скрытых слоев более трех не имеет смысла в силу теоретических соображений, большой ресурсоемкости обучения и опасности локализации нейронных сетей.

4. Количество нейронов в скрытых слоях. Определяется линейным перебором для каждого варианта нейрокалибровки. Методика перебора приведена в разделе 3.

5. Схема нейрокалибровки:

- только нейронная сеть (НС);
- линейная калибровка + нейронная сеть с минимизацией 3D-ошибки. Нейронная сеть обучается минимизировать ошибку линейной триангуляции [6];
- нейронная сеть + линейная калибровка с минимизацией 2D-ошибки. Нейронная сеть дает поправку к 2D-координатам изображения, с учетом которой затем проводится линейная триангуляция. Нейронная сеть используется для исправления дисторсионных искажений на снимках [6].

6. Функция активации:

- сигмоида. При этом данные приводятся к диапазону от 0 до 1;
- гиперболический тангенс. Данные приводятся к диапазону от -1 до 1.

7. Алгоритм обучения [7, 8]:

- Extended Delta-Bar-Delta. Один из вариантов алгоритма обратного распространения ошибки первого порядка;
- Levenberg-Marquardt. Алгоритм второго порядка.

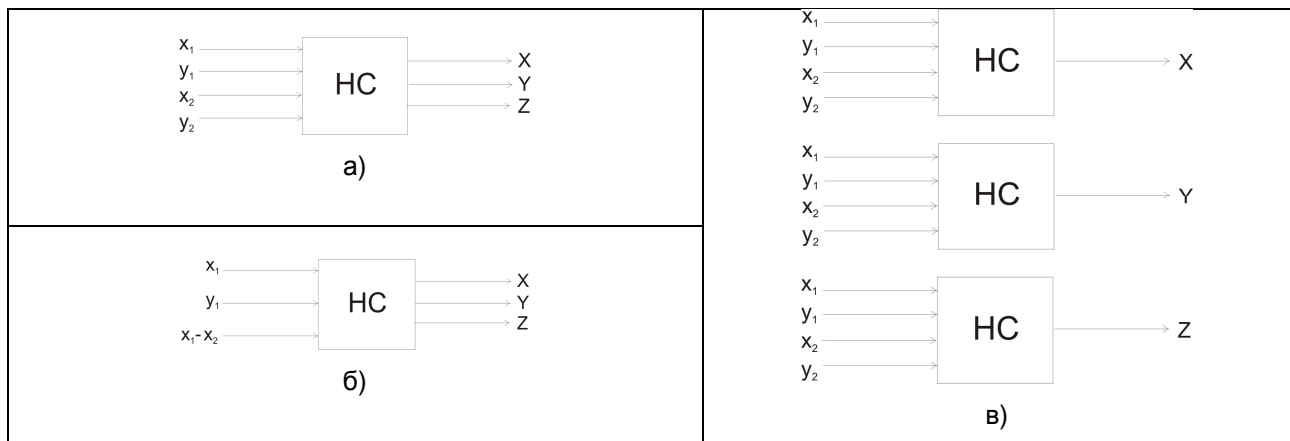


Рис. 2. Варианты использования нейронных сетей в задаче нейрокалибровки:

а) одномодульная сеть без препроцессинга: 4 входа, 3 выхода; б) одномодульная сеть с препроцессингом: 3 входа, 3 выхода; в) трехмодульная сеть без препроцессинга: три отдельные сети обучаются каждая на свою пространственную координату, 4 входа, 1 выход у каждой

В работе использовались следующие наборы данных:

- реальные данные для камер с небольшими нелинейными искажениями (1890 точек);
- синтетические данные с линейной моделью камер (2152 точек).

3. Методика нейрокалибровки

В каждом из приведенных выше методов нейрокалибровки возникала задача определения оптимального количества нейронов в скрытых слоях. Чтобы корректно сравнивать результаты разных методов, была необходима унифицированная методика решения этой задачи. Основу методики составляли автоматический линейный перебор и выбор лучшей сети. Все сети обучались методом SaveBest с алгоритмом обучения Extended-Delta-Bar-Delta (EDBD) и с использованием точек перехода (Transition Points). Действие точек перехода состоит в уменьшении коэффициента обучения и коэффициента момента обучения при достижении определенного числа показов векторов обучающей выборки, что позволяет избежать осциллирования нейронной сети вокруг минимума функции ошибки. Критерием выбора лучшей сети был минимум по худшей из ошибок на массивах train и test. Однако для двух и более скрытых слоев простой линейный перебор всех вариантов крайне ресурсоемок, поэтому архитектура нейросети определялась согласно следующей методике:

1. Проводился перебор вариантов в широком диапазоне с большим шагом (обычно около 5 нейронов). В результате грубо определялись сети с достаточным для решения задачи количеством нейронов. Специфика этого этапа состояла в том, что даже при избыточном количестве нейронов в скрытом слое ошибка на тестовом множестве была мала. Однако при дальнейшем обучении такие сети локализовались (что было видно по моменту останова метода SaveBest) и показывали плохие результаты. Поэтому для дальнейшего обучения выбирались наиболее простые сети из показавших хороший результат (вошедшие в пятерку лучших).

2. В полученной на первом этапе области проводился перебор вариантов с малым шагом (обычно два нейрона) для трех разных инициализаций матрицы весовых коэффициентов.

3. Для лучшей сети, полученной на втором шаге, проводилось обучение для 10 разных инициализаций матрицы весовых коэффициентов.

4. Лучшая сеть шага 2 и найденный на третьем шаге вектор инициализации использовались для обучения сети методом SaveBest до 15 миллионов векторов с тестом каждые 30000 векторов.

На первых трех шагах обучение проводилось в «черновом» режиме – до 1,6 миллионов векторов с 50 тестовыми точками. Основные параметры алгоритма обучения EDBD приведены в табл. 1.

Таблица 1. Основные параметры алгоритма обучения EDBD

Параметр	Значение
Начальная инициализация матрицы весовых коэффициентов	от $-0,3$ до $0,3$
Начальное значение коэффициента обучения	$0,3$
Начальное значение коэффициента момента	$0,3$
Точки перехода	$7E5, 2,1E6, 4,9E6, 10,5E6$ векторов
Максимальное значение коэффициента обучения	$0,2333$ (уменьшается вдвое при достижении точек перехода)
Максимальное значение коэффициента момента	$0,1667$ (уменьшается вдвое при достижении точек перехода)
Шаг изменения коэффициента обучения	$0,01$ (уменьшается вдвое при достижении точек перехода)
Шаг изменения коэффициента момента	$0,01$ (уменьшается вдвое при достижении точек перехода)

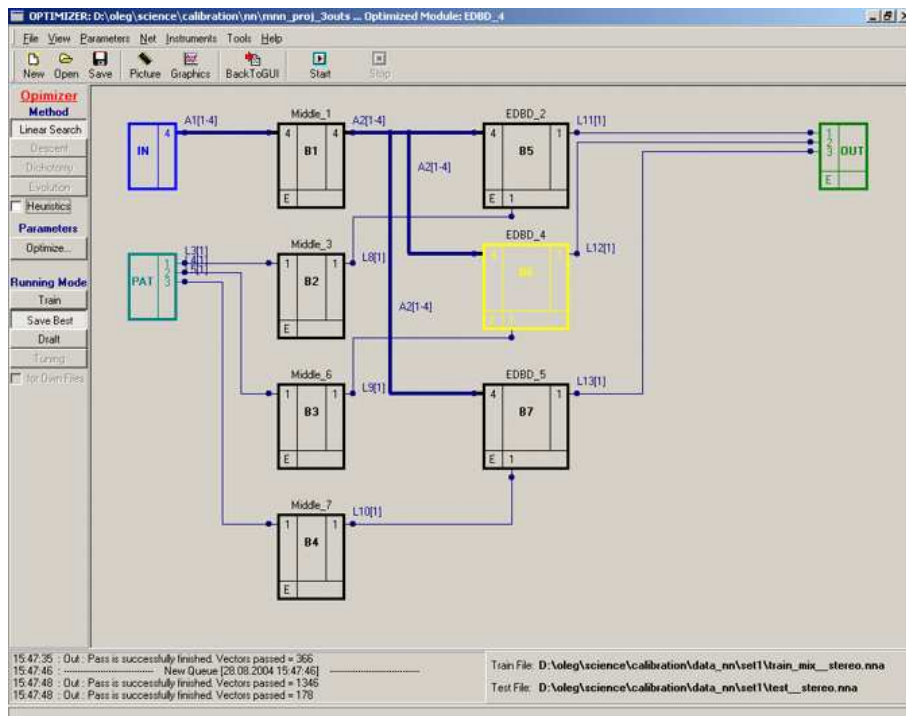


Рис. 3. Рабочее окно оптимизатора нейронных сетей программной системы MNN CAD

В работе использовалась программная система для работы с модульными нейронными сетями MNN CAD [9], включающая в себя систему автоматизированного проектирования модульных нейронных сетей, систему обучения модульных нейронных сетей и систему оптимизации архитектуры и параметров обучения нейронных сетей. Кроме этого, MNN CAD поддерживает широкий набор нейросетевых парадигм и включает разнообразные модули по предварительной обработке данных. Первые три шага в приведенной выше методике нейрокалибровки осуществлялись при помощи оптимизатора MNN CAD. Рабочее окно MNN CAD в режиме оптимизатора приведено на рис. 3.

4. Сравнение методов нейрокалибровки на реальных данных

Для калибровки использовались снимки калибровочного стенда [10] веб-камерами с малой дисторсией. Погрешность измерения 3D координат была порядка 5 мм, 2D координаты измерялись с точностью до пиксела. Данные были произвольно разбиты на три группы: train, test и valid (табл. 2) таким образом, чтобы точки, соответствующие одному положению стенда, попадали в одну группу (что в нашей постановке эксперимента позволяет верно оценить способность модели к обобщению и триангуляции в рабочем режиме). Обучающие векторы группы train были случайным образом перемешаны. Группы train и test использовались для обучения НС методом SaveBest, а также для калибровки аналитическими методами. Группа valid использовалась для проверки точности триангуляции на новых для метода данных.

Таблица 2. Характеристики реальных калибровочных данных

Группа	Количество точек	% от общего числа точек
Train	1346	71,2
Test	366	19,3
Valid	178	9,5
Всего	1890	100

В результате предварительных экспериментов были получены следующие общие результаты, повлиявшие на дальнейший ход исследования:

- обучение нейронных сетей алгоритмом Levenberg-Marquardt давало большие ошибки, чем обучение алгоритмом EDBD. Качественно это можно объяснить тем, что поверхность ошибки сильно «изрезана», поэтому сглаживание алгоритмов второго порядка не позволяет найти хороший локальный минимум. Это подтверждается также заметным уменьшением ошибки при прохождении точек перехода;

- пакетное обучение (off-line) оказалось заметно хуже, чем обучение на отдельных векторах (on-line). Это можно объяснить замедлением скорости схождения метода. Даже при on-line обучении требуется порядка 15 миллионов итераций, при пакетном обучении сеть за разумное время не доучивается;

- были найдены параметры алгоритма обучения EDBD (табл. 1), изменение которых ухудшало сходимость метода или приводило к незначительным частичным улучшениям лишь на

части рассматриваемых моделей. Эти параметры использовались во всех последующих экспериментах;

– смешанные модели (нейронная сеть + линейная калибровка) оказались нерабочими. Линейная калибровка основана на минимизации среднеквадратической ошибки и не учитывает ее знак. Поэтому даже для близких входных векторов знак ошибки практически случаен, хоть и близок по абсолютной величине. Обучение нейронных сетей на ошибку линейного метода в результате оказывается обучением на случайный шум, поэтому сеть не учится. Скорее всего, автор работы [0] использовал в качестве ошибки известную ему при генерировании данных дисторсионную поправку к линейному методу, однако при реальной калибровке эта информация недоступна.

Методы нейрокалибровки сравнивались со следующими аналитическими методами:

1. Линейный. Используется линейная (pin-hole) модель камеры [1], элементы проекционной матрицы ищутся методом псевдоинверсии [11].

2. Bouguet без дисторсии. Использована модель камеры Bouguet [3, 4]. При этом все коэффициенты дисторсии считались равными нулю и в калибровке не участвовали. Использовался метод калибровки Bouguet, доступный в Интернет в виде модулей Matlab [0]. Его суть заключается в нахождении начального приближения с помощью прямого линейного преобразования (Direct Linear Transformation, DLT) с последующим градиентным спуском.

3. Bouguet с k_1 . Использованы модель камеры Bouguet и калибровка Bouguet. При этом в калибровке участвовал только первый коэффициент радиальной дисторсии k_1 . Такая калибровка сходна с калибровкой по Tsai [2].

4. Bouguet с дисторсией. Здесь также использована модель камеры Bouguet, но с учетом всех пяти коэффициентов дисторсии (три радиальных и два тангенциальных). Метод калибровки тот же, что и в предыдущих двух пунктах.

Результаты работы нейросетевых и аналитических методов калибровки представлены в табл. 3. Для нейронных сетей конфигурация сети указана в виде $N_1-N_2-...-N_k$, где k – количество слоев (включая входной); N_k – количество нейронов в слое. В случае многомодульных сетей указаны конфигурации всех модулей. «Препроцессинг» обозначает описанный выше препроцессор, сокращающий количество входов до трех. Использование гиперболического тангенса в качестве функции активации обозначено как «th()». «Mean» означает среднюю ошибку триангуляции, «STD» – стандартное отклонение и «Max» – максимальную ошибку. Все результаты приведены в миллиметрах. Цветом выделены лучший нейросетевой и лучший аналитический методы.

Таблица 3. Ошибка триангуляции на реальных данных, мм

Метод	Train			Test			Valid		
	Mean	STD	Max	Mean	STD	Max	Mean	STD	Max
Линейный	7,60	5,51	38,6				7,05	6,03	29,8
Bouguet без дисторсии	7,58	6,11	44,1				8,61	7,57	33,5
Bouguet с k_1	7,48	6,46	39,3				6,94	5,02	23,4

Bouguet с дисторсией	7,16	5,89	39,8				6,37	5,80	27,3
НС, 1модуль, 4-23-3	18,4	12,2	71,5	20,0	12,6	54,4	31,1	29,2	130
НС, 1модуль, 4-19-13-3	6,74	5,26	37,9	8,33	5,56	31,4	9,92	7,47	39,4
НС, 1модуль, 4-19-13-7-3	6,20	4,79	33,3	8,05	5,35	26,5	8,59	6,95	35,9
НС, 1модуль, th(), 4-19-13-7-3	5,74	4,15	27,5	7,65	5,07	27,7	9,79	7,07	38,7
НС, 3 модуля X : 4-7-1, Y : 4-7-1, Z : 4-7-1	18,4	12,7	68,0	17,3	9,81	44,0	28,8	16,4	93,4
НС, 3 модуля X : 4-9-7-1, Y : 4-7-5-1, Z : 4-11-9-1	6,60	5,57	40,3	7,89	5,87	30,8	7,99	7,39	43,6
НС, 3 модуля, препроцессор X : 3-7-5-1, Y : 3-7-5-1, Z : 3-7-5-1	6,46	5,37	36,0	7,67	5,61	28,7	7,48	6,52	44,6
НС, 3 модуля, препроцессор, th(), X : 3-7-5-1, Y : 3-5-4-1, Z : 3-7-5-1	5,73	4,21	24,7	6,34	4,71	27,5	6,98	5,07	24,3
НС, 3 модуля, препроцессор X : 3-7-5-1, Y : 3-7-5-3-1, Y Z : 3-7-5-1	5,72	4,54	28,6	6,78	4,82	24,5	7,33	5,08	31,3
НС, 3 модуля, препроцессор, th(), X : 3-7-5-1, Y : 3-6-4-2-1, Z : 3-7-5-1	5,31	3,69	21,0	5,87	3,48	19,3	6,54	4,60	21,5
НС, 3 модуля, препроцессор, th(), X : 3-7-5-3-1, Y : 3-6-4-2-1, Z : 3-7-5-3-1	5,33	3,68	20,9	5,80	3,52	19,3	6,24	4,65	21,5

Согласно табл. 3, двухслойные и трехслойные нейронные сети показывают ошибку триангуляции, сравнимую с ошибкой аналитических методов, а трехслойная трехмодульная нейронная сеть с препроцессором и гиперболическим тангенсом в качестве функции активации даёт лучшие (примерно на 5% на массиве valid) результаты, чем метод Bouguet с учетом всех коэффициентов дисторсии. В целом, точность улучшается при увеличении числа скрытых слоев, использовании препроцессора и замене сигмоиды на гиперболический тангенс.

5. Сравнение методов нейрокалибровки на синтетических данных

Задачей данного этапа исследования была проверка выводов о выборе метода нейрокалибровки, сделанных на основании экспериментов на реальных данных. Исследовались такие вопросы:

- улучшает ли работу сети добавление второго и третьего скрытых слоев;
- улучшает ли работу метода обучение отдельной нейросети на каждую пространственную координату;
- улучшает ли работу сети замена сигмоидальной функции активации на гиперболический тангенс;
- улучшает ли работу сети препроцессинг;
- совпадает ли количество нейронов в скрытых слоях для каждого конкретного метода по сравнению с архитектурой, полученной на реальных данных.

Сравнение с аналитическими методами целью данного этапа не являлось по следующей причине. Данные генерировались при помощи той же модели камеры, которая использовалась для калибровки, поэтому аналитические методы калибровали стереопару практически идеально.

Данные были сгенерированы при помощи линейной модели камеры в масштабе, аналогичном массиву реальных данных. То есть абсолютные значения пространственных и 2D координат лежали в том же диапазоне, что и в случае реальных данных. Так, по глубине (Y) диапазон изменения пространственных координат составлял от 0 до 2000, по высоте (Z) – от 800 до 1700 и по ширине (X) – от 400 до 1600. Шаг сетки был равен 25 по всем координатам. Координаты на изображении принимали значения от 0 до 320 по x и от 0 до 240 по y , по аналогии с реальным сенсором. Затем данные были случайным образом разбиты на train, test и valid массивы. Характеристики синтетических данных приведены в табл. 4.

Таблица 4. Характеристики синтетических калибровочных данных

Группа	Количество точек	% от общего числа точек
Train	1499	69,7
Test	445	20,7
Valid	208	9,6
Всего	2152	100

Обучение и определение архитектуры нейросетей проводилось по той же методике, что и для реальных данных. Результаты калибровки на синтетических данных представлены в табл. 5. Цветом выделен лучший метод нейрокалибровки.

Таблица 5. Ошибка триангуляции на синтетических данных

Метод	Train			Test			Valid		
	Mean	STD	Max	Mean	STD	Max	Mean	STD	Max
1модуль, 4-21-3	35,4	16,5	131	37,5	16,6	123	38,5	15,3	131
1модуль, 4-19-13-3	11,4	5,46	43,2	11,9	5,66	29,4	12,2	5,72	37,0
1модуль, препроцессор, 3-21-3	13,6	5,94	32,2	14,6	5,99	29,8	13,4	6,15	28,8
1модуль, препроцессор, 3-21-15-3	5,54	2,73	19,3	5,96	2,96	20,8	5,51	2,93	18,1
1модуль, th(), 4-19-13-3	2,54	1,32	12,2	2,77	2,16	34,4	2,65	1,49	16,5
1модуль, препроцессор, th(), 3-21-15-3	2,15	1,22	8,83	2,24	1,25	10,1	2,30	1,29	6,98
1модуль, препроцессор, th(), 3-21-15-9-3	2,23	1,35	10,9	2,43	1,89	19,9	2,41	1,50	12,8
3 модуля, X : 4-7-1, Y : 4-3-1, Z : 4-5-1	34,8	16,0	79,5	36,0	15,5	78,8	37,6	15,2	74,0
3 модуля, X : 4-7-5-1, Y : 4-11-5-1, Z : 4-9-5-1	11,5	4,68	31,1	11,7	4,86	30,1	11,5	4,73	29,5
3 модуля, препроцессор, X : 3-9-5-1, Y : 3-9-5-1, Z : 3-7-5-1	3,56	2,04	14,0	3,76	2,18	14,5	3,49	2,14	15,4
3 модуля, препроцессор, th(), X : 3-9-5-1, Y : 3-7-5-1, Z : 3-7-5-1	1,32	0,74	6,11	1,29	0,77	6,01	1,28	0,67	4,59
3 модуля, препроцессор, th(), X : 3-7-5-3-1, Y : 3-7-5-3-1, Z : 3-7-5-3-1	0,92	0,58	5,41	0,91	0,56	5,29	0,92	0,50	2,98

Проанализируем полученные результаты:

– при прочих равных добавление второго скрытого слоя значительно улучшает точность триангуляции, добавление третьего скрытого слоя в трехмодульных сетях еще увеличивает точность;

– при прочих равных добавление препроцессора значительно улучшает точность;

– замена сигмоидальной функции активации на гиперболический тангенс повышает точность калибровки как при использовании препроцессора, так и без него;

– обучение отдельной сети на каждую координату незначительно улучшает точность по сравнению с одномодульной сетью в случае отсутствия препроцессинга и сигмоидальной функции активации (при этом, однако, значительно уменьшается максимальная ошибка); при использовании препроцессора трехмодульная сеть показывает значительно лучшие результаты, чем аналогичная ей одномодульная; при комбинировании препроцессора и гиперболического тангенса в качестве функции активации, точность еще улучшается;

– архитектура (количество нейронов в скрытых слоях) нейронных сетей практически точно совпадает для реальных и синтетических данных.

6. Заключение

В работе исследованы различные методы нейрокалибровки и показано, что они по точности не уступают аналитическим методам. Трехмодульная трехслойная архитектура ($X : 3-7-5-3-1$, $Y : 3-6-4-2-1$, $Z : 3-7-5-3-1$) с препроцессором и гиперболическим тангенсом в качестве функции активации показала даже немного (около 5%) лучшие результаты, чем лучший из рассмотренных аналитических методов, учитывающий три коэффициента радиальной и два коэффициента тангенциальной дисторсии [4]. Результаты были проверены на реальных данных, что позволяет говорить о применимости нейронных сетей в реальных задачах калибровки стереопары. Практически точное совпадение оптимального числа нейронов в скрытых слоях для синтетического и реального массивов данных позволяет использовать полученные архитектуры нейросетей как хорошее начальное приближение при калибровке стереопар, основанных на линейных камерах или камерах с малой дисторсией.

На основании проведенных исследований можно сформулировать следующие практические рекомендации по процедуре нейрокалибровки стереопар со слабыми нелинейными искажениями:

– нейронные сети с одним скрытым слоем показывают крайне плохую точность триангуляции. Точность значительно улучшается при переходе к сетям с двумя, а затем с тремя скрытыми слоями. Следует однако помнить, что увеличение числа скрытых слоев увеличивает опасность локализации сети и поэтому может не приводить к улучшению точности в рабочем режиме (поэтому нежелательно использование одномодульных трехслойных сетей);

– одномодульные двухслойные сети уже дают сравнимую с аналитическими методами точность триангуляции;

– при выборе числа нейронов в скрытых слоях можно пользоваться значениями, приведенными в табл. 3 и 5;

– использование препроцессора и гиперболического тангенса в качестве функции активации улучшают точность нейрокалибровки;

– для обучения нейронных сетей необходимо большое количество итераций. В нашем случае точность практически переставала улучшаться после 15 миллионов показов векторов обучающей выборки для параметров алгоритма обучения, приведенных в табл. 1. Также крайне желательно использование точек перехода, в которых уменьшаются коэффициенты скорости обучения, т.к. поверхность ошибки, видимо, сильно «изрезана».

Дальнейшего исследования требуют вопросы нейрокалибровки стереопары со значительной дисторсией, устойчивости нейрокалибровки к зашумленности данных и зависимости точности нейрокалибровки от количества калибровочных точек. Интересен также вопрос применения в данной задаче метода локальных экспертов на основе модульных нейросетей [0].

СПИСОК ЛИТЕРАТУРЫ

1. Armangue X., Salvi J., Battle J. A comparative review of camera calibrating methods with accuracy evaluation // Pattern Recognition. – 2002. – Vol. 35 (7). – P. 1617–1635.
2. Tsai R.Y. A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses // IEEE Int. Journal on Robotics and Automation. – 1987. – Vol. 3. – P. 323–344.
3. Heikkila J., Silven O. A Four-step Camera Calibration Procedure with Implicit Image Correction // CVPR97. – 1997. – P. 1106–1112.
4. Bouquet J.-Y. Camera Calibration Toolbox for Matlab // http://www.vision.caltech.edu/bouquetj/calib_doc.
5. Галушкин А.И. Нейрокомпьютеры и их применение. – М.: ИПРЖ, 2002. – Кн. 6: Нейроматематика. – 442 с.
6. Yongtae D. Application of Neural Networks for Stereo-Camera Calibration // International Joint Conference on Neural Networks. – 1999. – P. 2719–2722.
7. Haykin S. Neural Networks A Comprehensive Foundation // Englewood Cliffs. – 2nd ed. – NJ, Prentice-Hall, 1999.
8. Reed R.D., Marks R.J. Neural smithing: supervised learning in feedforward artificial neural networks. – Cambridge, MA: MIT Press, 1999. – 346 p.
9. Система автоматизированного проектирования модульных нейронных сетей CAD MNN / Резник А.М., Куссуль М.Э., Сычов А.С., Садовая Е.Г., Калина Е.А. // Математические машины и системы. – 2002. – № 3. – С. 28–36.
10. Киреев О.С. Автоматизированная калибровка стереопары в лабораторных условиях // Математичні машини і системи. – 2004. – № 1. – С. 86–100.
11. Алберт А. Регрессия, псевдоинверсия и рекуррентное оценивание. – М.: Наука, 1977. – 224 с.
12. Галинская А.А. Модульные нейронные сети: обзор современного состояния разработок // Математичні машини і системи. – 2003. – № 3, 4. – С. 87–102.