

УДК 004.622:007.52

А. Е. Покинтелица

ФГБНУ «Институт проблем искусственного интеллекта», г. Донецк,
283048, г. Донецк, ул. Артема, 118-б

ПРОБЛЕМЫ И СПЕЦИФИКА РЕДУКЦИИ ДАННЫХ В АВТОНОМНЫХ РОБОТОТЕХНИЧЕСКИХ СИСТЕМАХ

A. E. Pokintelitsa

FSBSI «Institute of Problems of Artificial intelligence», Donetsk
283048, Donetsk, Artema str., 118-b

PROBLEMS AND FEATURES OF DATA REDUCTION IN AUTONOMOUS ROBOTIC SYSTEMS

Вычислительные ресурсы мобильных робототехнических систем той или иной степени автономности весьма ограничены. В настоящей статье рассматривается проблема уменьшения объема и (или) упрощения структуры данных, обрабатываемых в ограниченных по мощности вычислительных системах. Выполнен анализ методов редукции данных и описаны их алгоритмы. Выделены ключевые понятия, необходимые для формализации метода редукции: целевой показатель редукции и критерий для процедуры сепарации данных по степени значимости. Исследованы характеристики процессов обработки информации, приводящие к редукции данных.

Ключевые слова: обработка данных, редукция данных, робототехнические системы.

Computing resources of mobile robotic systems of various autonomy levels are fairly limited. This paper considers the problem of reducing the amount and (or) simplifying the structure of data, processed in power-limited computing systems. The data reduction methods are analyzed and their algorithms are described. The key concepts for the formalization of a reduction method are identified: the reduction target indicator and the criterion for the data separation procedure by degree of their significance. The characteristics of information processing processes which result in data reduction are examined.

Key words: data processing, data reduction, robotic systems.

Введение

Стремительное развитие робототехники и технической кибернетики актуализирует изучение общих закономерностей получения, хранения, преобразования и передачи информации. Упорядоченная информация, представленная в формализованном виде и пригодная для дальнейшей обработки, – формирует данные. Способы представления информации в виде данных разрабатывались человеком и для нужд человека. Этот факт находит свое отражение при изучении структуры и форматов данных, применяемых для хранения и передачи различного рода информации. Однако, с точки зрения устройства, выполняющего обработку данных для достижения каких-либо целей, такое представление информации может быть избыточным. Данное обстоятельство порождает вопрос о возможности выявления ключевых компонентов в массивах данных, избирательная обработка которых позволит сократить нагрузку на вычислительные мощности устройства и при этом существенно не повлияет на эффективность выполнения робототехническим устройством своих функциональных задач.

Понятие редукции данных включает в себе как сокращение объема анализируемой информации, так и упрощение формы сложноструктурированных данных в результате некоторых преобразований. Дальнейшая обработка данных может включать использование средств и методов искусственного интеллекта в условиях ограниченных вычислительных ресурсов, что порождает проблему рационализации используемых алгоритмов – понижения их сложности и упрощения методов вычислений.

Настоящая работа направлена на изучение свойств данных как объекта – носителя информации и проводилась в рамках фундаментального исследования по разработке концепции контекстно-независимого подхода к выявлению ключевых компонентов в массивах сложноструктурированных данных для создания методов редукции входных данных.

Проблематика фундаментального исследования заявленного направления заключается в том, что для формирования концепции подхода к выявлению ключевых компонентов в поступающих на обработку данных сперва следует абстрагироваться от конкретного информационного содержания данных. Проследить, существуют ли взаимосвязи, закономерности, которые охарактеризуют данные как систему элементов, выполняющих определенную функцию – представлять содержащуюся в них информацию в упорядоченном виде, пригодном для анализа и дальнейшей обработки. Выявление таких общих связей послужит ключом к установлению способа взаимодействия компонентов данных как элементов некоторой системы и позволит определить структуру этой системы. Это, в свою очередь, сделает возможным построение моделей, в том числе и математических, для формального описания взаимосвязей свойств комплексов данных и характеристик, определяющих состояние, а также поведение робототехнической системы.

Специфика проводимого исследования состоит в том, что вопросы редукции данных рассматриваются в отношении управления робототехническими системами [1] и, в частности, управления автономными робототехническими системами (АРС). В связи с этим для решения проблем адаптации современных методов, применяемых для интеллектуальной обработки данных, к ограниченной производительности микропроцессоров, на базе которых функционируют АРС, необходимо исследовать возможность создания как можно менее ресурсозатратных методов редукции входных данных.

Изложенные выше суждения являются целеполагающей основой исследования и задают направленность данной работы. В качестве объекта исследования выбраны данные, поступающие на вход автономных робототехнических систем при взаимодействии с окружающей средой. Предметом исследования выступают методы редукции входных данных.

Современное состояние вопроса

Анализируя современное состояние проблемы, было замечено, что развитие сенсорной техники как основного источника данных для функционирования робототехнических устройств, стимулирует рост интереса к изучению способов организации информации и методов обработки сформированных данных [2-5]. Это является отражением связи массового потребления высокотехнологичной продукции с интенсивностью научных разработок в данной сфере. Однако анализ публикаций из открытых источников свидетельствует о том, что фундаментальных исследований это касается куда в меньшей степени.

Практически любая интеллектуальная робототехническая система содержит модуль управления на основе искусственной нейронной сети. Решение сложных задач средствами искусственного интеллекта требует использования высокопроизводительных компьютеров, объединенных в вычислительные сети, что делает невозможным их внедрение в малогабаритные АРС.

В настоящее время с целью аппаратной минимизации нейронной сети применяются следующие методы:

- 1) выбор оптимальной структуры нейронной сети;
- 2) выбор оптимального метода решения разностного или дифференциального уравнения нейронной сети;
- 3) отказ от обучения нейронной сети. В микропроцессор робота встраивается уже предварительно обученная нейронная сеть без возможности дальнейшего обучения;
- 4) использование переменной разрядности умножителей;
- 5) сокращение (редукция) объема входных и выходных данных.

Тем не менее, даже применение вышеуказанных мер в комплексе порой недостаточно, чтобы внедрить нейронную сеть в устройство, доступное широкому кругу потребителей.

Подавляющая масса используемых робототехнических устройств спроектированы с применением бионического (биомиметического) подхода, при котором принципиальные идеи или образы элементов технологического устройства заимствуются из живой природы. Нередко создатели роботов наделяют их антропоморфными чертами. Следствием этого является ярко выраженный акцент на применении систем машинного зрения как одного из основных способов получения информации о состоянии окружающей среды.

Технологии машинного зрения объединяют в себе знания из большого количества смежных отраслей. Сегодня в данной области можно определить основные направления научно-технического поиска – совершенствование систем технического зрения (СТЗ) и программно-алгоритмического обеспечения к ним, а также создание специализированных вычислительных систем [4].

Таким образом, теоретические исследования проблемы редукции данных в автономных робототехнических системах проводились, основываясь на анализе путей решения частных задач прикладного характера из данной и смежных областей.

Данные как способ представления информации в упорядоченном виде

Проблема редукции данных, обрабатываемых в ограниченных по мощности вычислительных системах, обусловлена тем, что алгоритмы редукции, по своей сути, являются эвристическими и жестко привязаны к выполнению конкретной практической задачи. Разработка подобных алгоритмов сопряжена с глубоким анализом информационного контекста данных, их структуры и механизмов генерации.

Автором выдвигается гипотеза о том, что данные, поступающие на вход робототехнической системы при взаимодействии с окружающей средой и влияющие на деятельность этой системы, могут быть редуцированы без значимого ущерба для эффективности выполнения робототехническим устройством своих функциональных задач.

Для проверки данной гипотезы необходимо иметь метод выявления ключевых компонентов (наиболее значимых, несущих в себе основную информацию) в массивах данных, метод сепарации компонентов данных по степени значимости и метод оценки эффективности работы робототехнической системы в ответ на редуцирование входных данных.

Как было отмечено выше, для формирования концепции подхода к выявлению ключевых компонентов в поступающих на обработку данных следует абстрагироваться от конкретного информационного содержания данных. Выделить взаимосвязи, характеризующие данные как систему элементов, выполняющую определенную функцию – представлять информацию в упорядоченном виде, пригодном для анализа и обработки.

Изложенное выше можно проиллюстрировать следующей схемой (рис. 1):

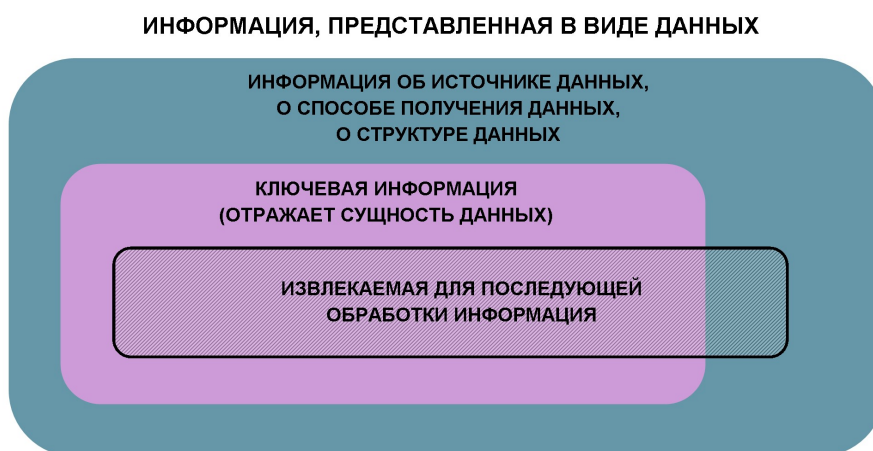


Рисунок 1 – Схематическое изображение информации, представленной в виде данных

Приведенная схема отражает и ту особенность, которая проявляется в практической невозможности строго разделить массив данных на «информативные» и «неинформативные». Положение и «ширина» границы раздела между этими двумя понятиями целиком и полностью зависит от выбранных методов сепарации.

Для большей ясности в вопросе о степени значимости компонентов данных, несущих в себе информацию различного рода, приведем ряд примеров. Так, в русском и многих других языках ключевая информация концентрируется вокруг согласных букв (звуков) [6], [7]. Вследствие чего, человеку, как носителю языка, не составит особого труда идентифицировать исходный текст по его образу, из которого были частично или полностью изъяты гласные буквы.

Принципы цифровой коммуникации накладывают свой отпечаток на способы представления хранимой и передаваемой информации. Передаваемые данные вмещают в себя информацию о методе их кодирования, об источнике, адреса «отправителя» и «получателя» и прочее. Различные протоколы, регламентирующие процедуру обмена данными, а также общеизвестные файловые форматы – стандартизированные формы кодирования информации, являются ярким тому примером.

Рассмотрим подробнее вопрос представления зрительной информации в виде файлов изображений. BMP (от англ. *Bitmap Picture*) – формат, разработанный для хранения однослойных растровых изображений. В общем случае информация о цвете хранится в виде двумерного массива, где цвет каждого пикселя задается в соответствии с цветовой моделью *RGB* комбинацией трех беззнаковых целых чисел. Глубина цвета изображения определяется количеством бит, необходимым для описания цвета. Так, при глубине цвета 8 бит цвет пикселя задают индексом из таблицы цветов, а при глубине цвета более 8 бит – непосредственным значением в цветовой модели *RGB*. Данные в формате BMP образуют три основных блока: заголовок *Bitmap file header*, заголовок *DIB header* и блок *Pixel array*, где хранятся данные о цвете пикселей.

Данные о цвете точек изображения, выраженные в математической форме, представляют собой аналог декартовых координат в n -мерном, а чаще трехмерном, пространстве. Существуют различные модели представления цвета, в основе которых лежит понятие базисных компонент (не обязательно цветовых). Во избежание путаницы отметим, что здесь и в двух следующих абзацах термин «компонента» имеет отношение к векторному представлению данных, в остальных же случаях понятие «компонент» трактуется как составная часть некоторого сложного объекта. Таким образом, комбинируя базисные компоненты можно воссоздать любой цвет из данного цветового пространства.

Одной из наиболее популярных цветовых моделей, используемых в компьютерной графике, является *RGB*. Для обработки видеосигналов чаще используют модель $YC_B C_R$. При переходе от *RGB* к $YC_B C_R$ информативная значимость данных смещается в сторону Y -компоненты – компоненты яркости [8]. Здесь Y -компонента выступает носителем ключевой информации с позиции зрительного ее восприятия. Если же нас интересует информация о цвете изображений, то нам следует сместить акцент в сторону цветоразностных составляющих – компонент C_B и C_R . Примечательно и то, что вопрос о критерии «значимости» решается автоматически с применением формул преобразования при переходе от *RGB* к $YC_B C_R$. Согласно рекомендации ITU-T T.871 (стандарт ISO/IEC 10918-5), формулы преобразования имеют следующий вид

$$\begin{aligned} Y &= \text{Min}(\text{Max}(0, \text{Round}(0,299 \cdot R + 0,587 \cdot G + 0,114 \cdot B)), 255), \\ C_B &= \text{Min}(\text{Max}(0, \text{Round}(-0,1687 \cdot R - 0,3313 \cdot G + 0,5 \cdot B + 128)), 255), \\ C_R &= \text{Min}(\text{Max}(0, \text{Round}(0,5 \cdot R - 0,4187 \cdot G - 0,0813 \cdot B + 128)), 255). \end{aligned}$$

Здесь компоненты Y , C_B и C_R должны быть преобразованы в целые числа из диапазона от 0 до 255. Значения в модели $YC_B C_R$ с 256 уровнями на компоненту могут быть вычислены непосредственно из значений цветовых каналов *RGB*, в которых черный цвет представлен комбинацией (0, 0, 0), а белый – комбинацией (255, 255, 255).

Говоря о матричном представлении, структурные характеристики данных в популярных форматах изображений отражают и тот факт, что представленная информация поступает от светочувствительных датчиков, расположенных на прямоугольной физической матрице, и является комплексированной на аппаратном уровне.

Таким образом, избирательная обработка значимых компонентов данных и отсеивание менее значимых компонентов позволят сократить нагрузку на вычислительные мощности устройств, выполняющих обработку этих данных.

Редукция данных в робототехнических системах

Специфика настоящей работы подразумевает, что все приведенные умозаключения и дальнейшие теоретические исследования вопросов редукции данных рассматриваются в отношении управления робототехническими системами и АРС в частности.

Данные – это та информация, которая поступает на вход робототехнической системы при взаимодействии с окружающей средой (окружением). В качестве источника данных могут выступать датчики, сенсоры, подключаемые модули и устройства (камеры, преобразователи сигналов и др.).

Автономные робототехнические системы – это системы, которые могут получать и самостоятельно обрабатывать информацию об окружающей среде и собственном состоянии, способны длительно работать без вмешательства оператора и самостоятельно выполнять свои функции. Внедрение элементов искусственного интеллекта в систему управления автономного робота наделяет его способностью обучаться в процессе работы и адаптироваться к изменениям в окружающей среде.

Для оценки влияния процедуры редуцирования входных данных на эффективность выполнения робототехническим устройством своих функциональных задач определим показатель редукции как меру отношения объемов данных до и после редукции. Конкретный вид (формула) для выражения этой меры будет зависеть от вида данных и метода их редукции.

Поведение робота можно оценить с помощью набора тестовых заданий. Подсчет соотношения успехов и неудач в серии испытаний будет служить методом оценки эффективности работы робототехнической системы в ответ на редуцирование входных данных.

Процессы обработки информации, приводящие к редукции данных

Задачи уменьшения объема и (или) упрощения структуры обрабатываемых данных требуют тщательного анализа свойств этих данных. В этой связи, применяемые решения зачастую относятся к эвристическим методам и строго ориентированы на конкретный тип задач. Однако, эвристика, в сравнении с точным алгоритмом, не гарантирует нахождение оптимального решения, может не найти заведомо существующее решение, а также не исключает того, что найденное решение при определенных условиях окажется неверным. Впрочем, это не уменьшает практической ценности применяемых решений, комплексный анализ которых может дать ценные знания об общих закономерностях в процессах обработки информации.

Изучая процессы обработки информации, приводящие к редукции данных, мы отвлекались от конкретного информационного содержания данных, считая его несущественным. Ряд примеров наглядно проиллюстрирует это.

Рассмотрим метод повышения быстродействия непараметрических классификаторов с помощью редукции входных данных в области анализа документальной информации. Методика и этапы ее разработки детально изложены в [9] и [10].

На начальном этапе формулируется целевой показатель редукции. Он учитывает требования по соотношению точности и быстродействия алгоритма, лежащего в основе редукции. Так, указанный метод, должен обеспечивать сокращение исходного объема данных без существенных потерь в эффективности их обработки с точки зрения результатов этой обработки. Поскольку работа сконцентрирована на исследовании непараметрических методов классификации, то оценка эффективности предложенных решений основана на ошибке классификации. Таким образом, метод редукции должен удовлетворять сформулированному целевому показателю.

На следующем этапе происходит сепарация данных, подлежащих редуцированию, методом, основанном на кластеризации данных, для чего вводится метрика в пространстве признаков рассматриваемых объектов. Массив объектов, носителей данных, разбивается на два класса, с позиции удовлетворения требованиям целевого показателя. Приводится также критерий, который лежит в основе процедуры разделения объектов на два класса.

После проверки объектов на возможность их объединения происходит редукция исходного массива.

Резюмируя, выделим следующие этапы редукции в рассмотренном примере:

- 1) задается целевой показатель редукции;
- 2) формулируется критерий для процедуры сепарации;
- 3) по выработанному алгоритму производится редукция исходных данных;
- 4) формируется новый массив данных.

Контроль эффективности происходит с помощью заранее подготовленных экзаменационных данных.

В приведенном выше примере метрикой в пространстве признаков рассматриваемых объектов является евклидово расстояние. Однако, при увеличении размерности признакового пространства вычислительная сложность алгоритмов, основанных на определении евклидова расстояния, растет экспоненциально. В работе [11] описан альтернативный подход, который позволяет обойти указанный недостаток. В результате был разработан метод непараметрической кластеризации данных, основанный на гистограммном анализе образов в признаковом пространстве.

Алгоритм редукционной кластеризации в описанном методе включает следующие этапы:

- 1) задается начальный пороговый уровень – аналог целевого показателя редукции;
- 2) для каждого признака выполняется построение координатных профилей – проекций многомерной плотности точек на координатную ось признакового пространства;
- 3) выделяются кластерные зоны – области, в которых плотность точек превышает установленный пороговый уровень (число построенных зон будет являться оценкой числа кластеров);
- 4) проводится процедура локализация кластерных зон – оптимизация положения порогового уровня по числу кластерных зон и суммарному объему их покрытий по точкам признакового пространства;
- 5) выполняется редукция данных.

На финальном этапе по выбранным значениям порогового уровня заново строятся кластерные зоны профилей и размечаются индивидуальными локальными метками. Для разметки используются так называемые профильные векторы принадлежности

точек. Эти векторы определяют принадлежность точек бинам гистограммы каждого координатного профиля и находятся на этапе построения координатных профилей. Таким образом, для каждого координатного профиля все точки признакового пространства класса отмечаются метками кластерных зон и, как следствие, может быть построена кластерная структура полного признакового пространства.

С целью сокращения числа переменных в процессе предварительной обработки данных применяют также методы факторного анализа. В таком ключе факторный анализ можно рассматривать как метод редукции данных. Помимо прочего он дает представление о структуре взаимосвязей между переменными [12].

Анализ применяемых на практике методов редукции данных и их алгоритмов позволил установить общие черты рассматриваемых методов, выделить ключевые понятия, необходимые для формализации метода редукции: целевой показатель редукции и критерий для процедуры сепарации данных по степени значимости. Их определение является основополагающим для разработки алгоритма, по которому производится редукция исходных данных.

На основании анализа процессов обработки информации, результатом действия которых является уменьшение объема и (или) упрощение структуры обрабатываемых данных, выделены три основных процесса, приводящих к редукции данных: прореживание, агрегирование и трансформация.

Прореживание – процесс исключения части данных из набора или потока. Имеет место в системах реального времени, когда новые данные от датчика или группы датчиков регистрируются быстрее, чем система способна их обработать, не вызывая при этом задержку прогнозируемой реакции.

Агрегирование – процесс объединения сходных по структуре или типу данных от нескольких источников в один набор или поток с полным или частичным сохранением свойств данных. Частным случаем агрегирования можно считать комплексирование датчиков – процесс получения обобщенной и улучшенной информации с помощью интеграции информации от нескольких датчиков.

Трансформация – процесс преобразования данных от одного или нескольких источников в один набор или поток без сохранения первоначальных свойств данных.

При рассмотрении процессов обработки зрительной информации было замечено, что подавляющее количество используемых в настоящее время классификаторов изображений характеризуются антропоморфностью. Иначе говоря, принцип работы нейронных сетей копирует предполагаемый механизм обработки визуальной информации в человеческом мозге: сначала, в первом слое, расположен фильтр границ; затем, в следующем слое, – фильтр рисок с разными наклонами; далее идет слой, выделяющий очертания предметов (круг, квадрат, и т. п.); и так далее до конечного слоя, который относит объект к конкретному классу.

Альтернативный подход применяется в методе, основанном на использовании дискретного косинусного преобразования (ДКП) [13] для нейронных сетей, ориентированных на обработку изображений [8]. Характеризуется он тем, что для обычных изображений позволяет сократить количество операций умножения более чем в три раза. При приемлемом ухудшении четкости на изображении выигрыш может достигать до 10 раз. Кроме того, время обучения нейронной сети сокращается на 15-25%. При этом вероятность ложной классификации остается на прежнем уровне.

Тем самым показано, что классифицировать изображение можно в совершенно ином, «ортогональном» пространстве, полученном преобразованием Фурье или ДКП. Данный прием позволяет отнести этот метод к группе методов редукции данных на основе их трансформации. Опираясь на достоинства предложенного метода и его

ориентированность, можно утверждать, что применение данного метода будет особенно эффективно в приложении к робототехническим системам, снабженным системами машинного зрения.

Выводы

Проблема разработки унифицированного методологического аппарата для решения задач по уменьшению объема и (или) упрощению структуры обрабатываемых данных не теряет своей актуальности. Основные результаты, полученные в ходе теоретических исследований:

1. Выполнен анализ применяемых на практике методов редукции данных и описаны их алгоритмы. Установлены общие черты рассматриваемых методов. Выделены ключевые понятия, необходимые для формализации метода редукции: целевой показатель редукции и критерий для процедуры сепарации данных по степени значимости.

2. Исследованы характеристики процессов обработки информации, результатом действия которых является уменьшение объема и (или) упрощение структуры обрабатываемых данных. Выделены три основных процесса, приводящих к редукции данных: прореживание, агрегирование и трансформация.

Таким образом, изучение общих закономерностей процессов, приводящих к редукции данных, послужит базой для построения математической модели взаимосвязей входных и выходных характеристик в задачах управления автономными робототехническими системами.

Список литературы

1. Зуев В. М. Моделирование робота для инспектирования инженерных коммуникаций [Текст] / В. М. Зуев, О. А. Бутов, С. И. Уланов, А. А. Никитина // Проблемы искусственного интеллекта. – 2021. – № 4 (23). – С. 46-57.
2. Linåker F. Unsupervised on-line data reduction for memorisation and learning in mobile robotics. Ph. D. thesis [Текст] / Linåker F. Sheffield, 2003. – 194 (205) p.
3. Уланов А. А. Современное состояние и перспективы развития рынка робототехники в мире и России. : дис. ... канд. экон. наук: 08.00.14 [Текст] / Уланов А. А. – МГИМО (у) МИД РФ, Москва, 2021. – 173 с.
4. Технологии робототехники. – Текст : электронный // Агентство промышленного развития Москвы : официальный сайт. – Москва, 2020. – URL: <https://apr.moscow/content/data/3/05%20%D0%A2%D0%B5%D1%85%D0%BD%D0%BE%D0%BB%D0%BE%D0%B3%D0%B8%D0%B8%20%D1%80%D0%BE%D0%B1%D0%BE%D1%82%D0%BE%D1%82%D0%B5%D1%85%D0%BD%D0%B8%D0%BA%D0%B8.pdf> (дата обращения: 24.02.2023).
5. Машинное зрение и технологии сенсорики. – Текст : электронный // Агентство промышленного развития Москвы : официальный сайт. – Москва, 2020. – URL: <https://investmoscow.ru/media/3341142/%D0%BC%D0%B0%D1%88%D0%B8%D0%BD%D0%BD%D0%BE%D0%B5-%D0%B7%D1%80%D0%B5%D0%BD%D0%B8%D0%B5-%D0%B8-%D1%82%D0%B5%D1%85%D0%BD%D0%BE%D0%BB%D0%BE%D0%B3%D0%B8%D0%B8-%D1%81%D0%B5%D0%BD%D1%81%D0%BE%D1%80%D0%B8%D0%BA%D0%B8.pdf> (дата обращения: 24.02.2023).
6. Колоколов А. С. Исследование информативных признаков речевого сигнала [Текст] / А. С. Колоколов, И. А. Любинский, А. Ю. Мещеряков, В. П. Яхно // Научные технологии. – 2012. – Т. 13, № 2. – С. 041-047.
7. Арутюнян В. Г. Визуальное распознавание слов в условии нехватки информации: регистрация движений глаз (экспериментальное исследование на материале русского языка) [Текст] / В. Г. Арутюнян // Вопросы когнитивной лингвистики. – 2017. – № 4 (53). – С. 112-122.

8. Близно М. В. Распознавание изображений по коэффициентам дискретного косинусного преобразования [Текст] / М. В. Близно, В. М. Зуев, С. Б. Иванова // Искусственный интеллект: теоретические аспекты и практическое применение : материалы Донецкого международного круглого стола. – Донецк : ГУ «Институт проблем искусственного интеллекта» (ГУ «ИПИИ»), 2022. – 172 с. – С. 86-90.
9. Бородкин А. А. Разработка метода повышения быстродействия непараметрических классификаторов библиографических текстовых документов. : автореф... дис. кан. техн. наук: 05.13.01 [Текст] / А. А. Бородкин. – Москва : МЭИ, 2012. – 19 с.
10. Бородкин А. А. Методы повышения быстродействия непараметрических классификаторов в задачах обработки и анализа библиографических текстовых документов [Текст] / А. А. Бородкин, В. О. Толчеев, И. В. Некрасов // Приложение к журналу «Информационные технологии» №11. — 2013. – Издательство «Новые технологии», 30 с.
11. Дорогов А. Ю. Редукционная кластеризация высокоразмерных данных [Текст] / А. Ю. Дорогов // IV Междунар. науч. конф. по проблемам управления в технических системах (ПУТС-2021). Сб. докладов. Санкт-Петербург; 21-23 сентября, 2021 г. – СПб.: СПбГЭТУ «ЛЭТИ». 2021. – С. 248-251.
12. Голубков А. М. Применение метода каскадной редукции к решению задачи распознавания лиц [Текст] / А. М. Голубков, Д. М. Клионский // Известия СПбГЭТУ «ЛЭТИ». – 2019. – №8. – С. 47-53.
13. Чернов В. М. Быстрые алгоритмы дискретных косинусных преобразований коротких длин с минимальной вычислительной сложностью [Текст] / В. М. Чернов, М. А. Чичева // Известия Самарского научного центра Российской академии наук. – 1999. – Т. 1, № 2. – С. 241-248.

References

1. Zuyev V. M., Butov O. A., Ulanov S. I., Nikitina A. A. Modelirovanie robota dlia inspektirovaniia inzhenernykh kommunikatsii [Simulation of a robot for inspection engineering communications]. *Problemy iskusstvennogo intellekta* [Problems of artificial intellect], 2021, no. 4 (23), pp. 46-57.
2. Linåker F. *Unsupervised on-line data reduction for memorisation and learning in mobile robotics*. Ph. D. thesis. Sheffield, 2003, 194 (205) p.
3. Ulanov A. A. *Sovremennoe sostoianie i perspektivy razvitiia rynka robototekhniki v mire i Rossii*. Ph. D. thesis. Moscow, 2021, 173 p.
4. Tekhnologii robototekhniki, available at: <https://apr.moscow/content/data/3/05%20%D0%A2%D0%B5%D1%85%D0%BD%D0%BE%D0%BB%D0%BE%D0%B3%D0%B8%D0%B8%20%D1%80%D0%BE%D0%B1%D0%BE%D1%82%D0%BE%D1%82%D0%B5%D1%85%D0%BD%D0%B8%D0%BA%D0%B8.pdf> (accessed 24 February 2023).
5. Mashinnoe zrenie i tekhnologii sensoriki, available at: URL: <https://investmoscow.ru/media/3341142/%D0%BC%D0%B0%D1%88%D0%B8%D0%BD%D0%BD%D0%BE%D0%B5-%D0%B7%D1%80%D0%B5%D0%BD%D0%B8%D0%B5-%D0%B8-%D1%82%D0%B5%D1%85%D0%BD%D0%BE%D0%BB%D0%BE%D0%B3%D0%B8%D0%B8-%D1%81%D0%B5%D0%BD%D1%81%D0%BE%D1%80%D0%B8%D0%BA%D0%B8.pdf> (accessed 24 February 2023).
6. Kolokolov A. S., Lubinski I. A., Mestcherakov A. Yu., Yachno V. P. Issledovanie informativnykh priznakov rechevogo signala. *Journal Science Intensive Technologies*, 2012, vol. 13, no 3, pp. 041-047.
7. Arutunian V. G. Vizual'noe raspoznavanie slov v uslovii nekhvatki informatsii: registratsiia dvizhenii glaz (eksperimental'noe issledovanie na materiale russkogo iazyka). *Issues of Cognitive Linguistics (Voprosy Kognitivnoy Lingvistiki)*, 2017, no. 4, pp. 112-122.
8. Близно М. В., Зуев В. М., Иванова С. Б. Распознавание изображений по коэффициентам дискретного косинусного преобразования. *Iskusstvennyi intellekt: teoreticheskie aspekty i prakticheskoe primenenie. Materialy Donetskogo mezhdunarodnogo kruglogo stola*. Donetsk, SI IPAI, 2022, pp. 86-90.
9. BoroDkin A. A. Razrabotka metoda povysheniia bystrodeistviia neparametricheskikh klassifikatorov bibliograficheskikh tekstovykh dokumentov. Abstract of Ph. D. thesis. Moscow, 2012, 19 p.
10. BoroDkin A. A., Tolcheev V. O., Nekrasov I. V. *Metody povysheniia bystrodeistviia neparametricheskikh klassifikatorov v zadachakh obrabotki i analiza bibliograficheskikh tekstovykh dokumentov. Prilozhenie k zhurnalu «Informatsionnye tekhnologii»*, 2013, no. 11, 30 p.
11. Dorogov A. Yu. Reduktsionnaia klasterizatsiia vysokorazmernykh dannykh. *IV Mezhdunarodnaia nauchnaia konferentsiia po problemam upravleniia v tekhnicheskikh sistemakh. Sbornik dokladov*. Saint Petersburg, SPbETU "LETI", 2021, pp. 248-251.
12. Golubkov A. M., Kliionskii D. M. Primenenie metoda kaskadnoi reduksii k resheniiu zadachi raspoznavaniia lits. *Proceedings of Saint Petersburg Electrotechnical University*, 2019, no. 8, pp. 47-53.
13. Chernov V. M., Chicheva M. A. Bystrye algoritmy diskretnykh kosinusnykh preobrazovaniia korotkikh dlin s minimal'noi vychislitel'noi slozhnost'iu. *Izvestia of Samara Scientific Center of the Russian Academy of Sciences*, 1999, vol. 1, no. 2, pp. 241-248.

RESUME

A. E. Pokintelitsa

Problems and Features of Data Reduction in Autonomous Robotic Systems

The concept of data reduction implies a reduction in the amount of information and (or) simplification of a complex data structure as a result of certain transformations. This article considers the problem of data reduction and increasing the efficiency of data processing methods in microprocessor devices that run autonomous robotic systems.

The theoretical study of the issue mentioned above was carried out based on the analysis of specific applied problem solutions from this and related subject areas.

The analysis of data reduction methods used in practice is carried out and their algorithms are described. The general features of the considered methods are found out. The key concepts for the formalization of a reduction method are identified: the reduction target indicator and the criterion for the data separation procedure by degree of their significance.

The characteristics of data handling processes, that entail a reduction in data volume and (or) simplification of the structure of the processed data, are studied. Three main processes leading to data reduction are highlighted.

A study of the general patterns of processes leading to data reduction will make it possible to better understand the relationship between input and output characteristics in the tasks of controlling autonomous robotic systems, and will become the basis for the development of mathematical models of control systems.

РЕЗЮМЕ

А. Е. Покинтелица

Проблемы и специфика редукции данных в автономных робототехнических системах

Понятие редукции данных включает в себе сокращение объема информации и (или) упрощение формы сложноструктурированных данных в результате некоторых преобразований. В данной статье рассмотрена проблема редукции данных и повышения эффективности методов их обработки в микропроцессорных устройствах, на базе которых функционируют автономные робототехнические системы.

Теоретические исследования обозначенной проблемы проводились, основываясь на анализе путей решения частных задач прикладного характера из данной и смежных областей.

Выполнен анализ применяемых на практике методов редукции данных и описаны их алгоритмы. Установлены общие черты рассматриваемых методов. Выделены ключевые понятия, необходимые для формализации метода редукции: целевой показатель редукции и критерий для процедуры сепарации данных по степени значимости.

Исследованы характеристики процессов обработки данных, результатом действия которых является уменьшение объема и (или) упрощение структуры обрабатываемых данных. Выделены три основных процесса, приводящих к редукции данных.

Изучение общих закономерностей процессов, приводящих к редукции данных, позволит лучше понять взаимосвязи между входными и выходными характеристиками в задачах управления автономными робототехническими системами и послужит базой для построения математических моделей систем управления.

Статья поступила в редакцию 25.01.2023.