

КОГНИТИВНОЕ МАШИННОЕ ЗРЕНИЕ НА ОСНОВЕ ПАРАМЕТРИЧЕСКОГО АНАЛИЗА СТРУКТУРНЫХ ПРИМИТИВОВ ИЗОБРАЖЕНИЙ

СПИСОК ЛИТЕРАТУРЫ

1. PGI Accelerator Programming Model for Fortran & C [электронный ресурс] / The Portland Group, 2010 – Режим доступа: http://www.pgroup.com/lit/whitepapers/pgi_accel_prog_model_1.3.pdf
2. Romain Dolbeau. HMPP: A Hybrid Multi-core Parallel Programming Environment / Romain Dolbeau, Stephane Bihan, Francois Bodin – CAPS entreprise, 2007.
3. Nyland, L. Fast N-Body simulation with CUDA / L.Nyland, M. Harris // GPU Gems 3 – 2007 – с. 677.
4. CUDA C Best Practices Guide [электронный ресурс] / NVidia corporation, version 3.2 – с. 50 – режим доступа: <http://docs.nvidia.com/cuda/cuda-c-best-practices-guide/>
5. Кривов, М.А. Портируем на GPU и оптимизируем под CPU / М.А. Кривов, А.М. Казеннов // Журнал «Суперкомпьютеры», Весна 2011 – с. 43-45.
6. Кривов, М.А. Опыт портирования среды для HDR-обработки изображений на GPU и APU. / М.А. Кривов, М.Н. Притула, С.Г. Елизаров // Параллельные вычислительные технологии (ПаВТ'2012): труды международной научной конференции (Новосибирск, 26 – 30 марта 2012 г.), Челябинск: Издательский центр ЮУрГУ, 2012 – 774 с.
7. Кривов, М.А., Сине-зелено-красная OpenCL/ М.А. Кривов// Журнал «Суперкомпьютеры», Осень 2011 - с. 47-50.
8. Timothy, G. Mattson. Patterns for Parallel Programming / G. Mattson Timothy; Beverly A. Sanders; Berna L. Massingill – Addison-Wesley Professional, NY: 2009.
9. Janhunen, P. A positive conservative method for magnetohydrodynamics based on HLL and Roe methods / P. Janhunen – J. Comp. Phys., 1999.

Аспирант С.А. Гризан, д.т.н., проф., зав. каф. А.И. Легалов - каф. вычислительной техники Сибирского федерального университета

УДК: 004.81:159.9

УДК: 004.93'1; 004.932

УДК: 615.471; 681.32(075)

КОГНИТИВНОЕ МАШИННОЕ ЗРЕНИЕ НА ОСНОВЕ ПАРАМЕТРИЧЕСКОГО АНАЛИЗА СТРУКТУРНЫХ ПРИМИТИВОВ ИЗОБРАЖЕНИЙ

К.Б. Саниев

Проблема автоматического распознавания изображений ставится в общем виде как информационная задача извлечения из видеосигнала данных об априорно неопределённых сценах и объектах. Решение строится на операциях обнаружения визуальных структурных примитивов, определении достаточного набора их признаков, безэталонной классификации по характеристическим признаковым гистограммам и различению по форме. Приводятся примеры апробации элементов когнитивной технологии машинного зрения при распознавании изображений объектов различных классов и динамических сцен.

Ключевые слова: когнитивное машинное зрение, визуальные примитивы, безэталонная классификация.

Введение

Практические задачи автоматизации и информатизации производств, процессов и научных исследований требуют повышения эффективности всех видов систем технического зрения (СТЗ). Одной из основных целей этих разработок по-прежнему остаётся воспроизведение в распознающих автоматах информационных функций биологического зрительного восприятия. Качественное сравнение его информативности и методов автоматического анализа сигналов изображений показывает следующее.

- Современные технологии автоматического распознавания изображений, исходящие из методологии «распознавания обра-

зов», реализуют [1-3] частную информационную функцию классификации в виде проверки гипотез о наличии заранее заданных объектов. Априорная неопределённость множества (классов) объектов, которые могут быть различимы и опознаны, снимается частично посредством «ручного» задания M эталонных описаний, построенных на N эвристически сформированных признаках. Возможность расширения и коррекции множества эталонов в автоматическом режиме отсутствует. Информационная эффективность функции распознавания в этом случае принципиально ограничена количеством M заданных эталонов.

РАЗДЕЛ 2. ОБРАБОТКА СИГНАЛОВ И ДАННЫХ

- Информационная функция биологического восприятия обеспечивает организму наблюдение всех составляющих сцену объектов, оценку их биологической значимости для животного и управление поведением. При этом именно в процессе функционирования осуществляется расширение и коррекция множества распознанных (классов) объектов. Информационная эффективность такой функции распознавания ничем не ограничивается и оценивается количеством $K(t)$ выявляемых (классов) объектов ($K(t) \gg M$).

В связи с этим и задачу автоматического распознавания изображений предлагается ставить в общем виде как задачу извлечения из анализируемого сигнала данных о всех объектах наблюдаемой сцены.

Для определённости и отличия от «распознавания», будем называть более соответствующую биологическому восприятию информационную функцию «интерпретацией», а реализующий её алгоритм AL^{Reg} , преобразующий изображение I^k априорно неопределённой сцены в её формализованное описание D^Ω , $AL^{Reg}: I^k \xrightarrow{AL^{Reg}} D^\Omega$ считать необходимым ядром когнитивного машинного зрения и ассоциировать с «когнитивным видеоавтоматом» A^S

$$A^S = (\mathfrak{I}(\bar{x}, t), \Theta, D^\Omega, f1, f2).$$

Когнитивный видеоавтомат

Синтез когнитивного видеоавтомата A^S инициализируется утверждением (которое сформулировал ещё Д. Марр [4] в своём информационном подходе к зрению) о том, что существует всё-таки универсальное единое представление изображений множеством первичных элементов, которые естественно назвать «визуальными примитивами». Согласно физике формирования изображений, ими являются проекции отражающих/излучающих поверхностей реальных объектов наблюдаемой сцены. На плоском изображении визуальные примитивы представляют собой компактные подмножества пикселей с вычислимыми яркостно-геометрическими параметрами. Поэтому при определении [5-7] собственных структурных примитивов изображений как пространственно и амплитудно связанных подмножеств \mathcal{G}_j пикселей выявленное на изображении $\mathfrak{I}(\bar{x}, t)$ множество $\Theta = \{\mathcal{G}_j\}$ собственных примитивов удовлетворяет условиям $\mathcal{G}_i \cap \mathcal{G}_j = \emptyset$ одно-

значности и $\bigcup \mathcal{G}_j = \mathfrak{I}(\bar{x}, t)$ полноты. Такое множество визуальных примитивов содержит всю информацию о наблюдаемой сцене и составляющих её объектах, поскольку изображения объектов всегда будут представлены группами Θ^Ω примитивов, а наблюдаемые свойства $P_Q: \{p^{\Omega_1}, \dots, p^{\Omega_v}\}$ объектов могут быть составлены из характеристик $F_g: \{f^{g_1}, \dots, f^{g_\mu}\}$ составляющих их визуальных примитивов \mathcal{G}_j .

Соответственно, логика информационного анализа сигнала, выполняемого когнитивным видеоавтоматом, строится на параметрическом анализе множества визуальных примитивов.

По сравнению с «конечным автоматом» в состав когнитивного видеоавтомата A^S должны входить входной сигнал изображений $\mathfrak{I} = \mathfrak{I}(\bar{x}, t)$, множество $\Theta = \{\mathcal{G}_j(\bar{x}, F_g)\}$ визуальных примитивов, составляющих сигнал \mathfrak{I} и являющихся алфавитом для построения описаний выявляемых объектов, и функционал $f1: \mathfrak{I} \times f1 \rightarrow \Theta$ преобразования входного сигнала \mathfrak{I} в множество Θ примитивов. Полное определение когнитивного видеоавтомата $A^S = (\mathfrak{I}(\bar{x}, t), \Theta, D^\Omega, f1, f2)$ включает в себя также выходное множество $D^\Omega = \{\Omega_k, P^\Omega\}$ описаний (классов) объектов и функционал $f2: A \times f2 \rightarrow D^\Omega$ формирования на множестве примитивов выходного множества D^Ω .

Когнитивный видеоавтомат A^S обеспечивает преобразование $A^S: \mathfrak{I}(\bar{x}, t) \rightarrow D^\Omega$ видеосигнала $\mathfrak{I} = \mathfrak{I}(\bar{x}, t)$ в структурированный на классы D_{Cl}^Ω список D^Ω объектов и их характеристик $P^\Omega: \{p_1, \dots, p_\mu\}$ следующей последовательностью операций: обнаружение визуальных примитивов $\mathcal{G}_j(\bar{x})$; вычисление характеристик $F_g: \{f^{g_1}, \dots, f^{g_\mu}\}$ визуальных примитивов; классификация множества визуальных примитивов $\Theta = \{\mathcal{G}_j\}$ на «объектные» подмножества Ω_k , вычисление характеристик $P_Q: \{p^{\Omega_1}, \dots, p^{\Omega_v}\}$ объектов; классификация множества Ω_k на классы D_{Cl}^Ω .

Параметры визуальных примитивов

В общем случае описание произвольной сцены содержит данные о наблюдаемых в сцене объектах, их взаимном расположении и внешнем виде, а также их поведенческой активности. Эти данные представлены в видеосигнале статическими и динамическими параметрами, которые могут быть выявлены при анализе отдельных видеокадров и межкадровых изменений.

Множество P_g статических параметров включает в себя подмножества P_g^{br} , P_g^{gm} и P_g^{pl} , $P_g = P_g^{br} \cup P_g^{gm} \cup P_g^{pl}$ соответственно, амплитудных (яркостных), геометрических и топологических характеристик примитивов $\mathcal{G}_j(\tilde{x}, \tilde{y})$. В подмножестве P_g^{gm} геометрических параметров-признаков визуальных объектов наибольшую информативность имеет параметр формы, отыскание оптимальной формализации которого представляет отдельную проблему. Здесь перспективным решением представляется оценка двух характерных свойств силуэта – распределения внутренних областей («дыр») и конфигурации внешних «выступов».

Топологические параметры-признаки визуальных объектов, составляющие подмножество P_g^{pl} , предназначаются для выявления групп примитивов, соответствующих изображениям сложных объектов. Примером признака «объекта-группы» может быть меньшее расстояние между примитивами одной группы, чем между примитивами разных групп. Другими признаками «объекта-группы» могут служить отношения подобия и симметрии между примитивами.

Для единообразия рассмотрения статических и динамических свойств объектов введём дополнительно к понятию статических структурных примитивов $\mathcal{G}_j(\tilde{x}, \tilde{y})$, которые детектируются на одном – статическом – изображении, понятие динамических структурных примитивов [6], которые детектируются на сигнале видеоактивности. Видеоактивность ΔI представляет собой разницу соседних во времени изображений: $\Delta I(x, y, \Delta t) = I_2(x, y, t_2) - I_1(x, y, t_1)$.

Динамические примитивы $\partial \mathcal{G}_j(\tilde{x}, \tilde{y})$ видеоактивности $\Delta I(x, y, \Delta t)$ всегда соответствуют изменениям состояния объектов и содержат информацию об их поведении. Поэтому динамические признаки, составляющие

подмножество $P_{\partial \mathcal{G}}$, позволяют единообразно описывать любые ситуации событиями вида: появление объектов, исчезновение объектов, мерцание и движение объектов. Данные о «форме поведения» объектов содержатся во множестве $Tr^{\partial \mathcal{G}}: \{tr^{\partial \mathcal{G}1}, \dots, tr^{\partial \mathcal{G}k}, \dots, tr^{\partial \mathcal{G}m}\}$ траекторий, образуемых динамическими примитивами, и в параметрах этих траекторий. Основной проблемой при выявлении траекторий $tr^{\partial \mathcal{G}k}$ является отождествление разных динамических примитивов, детектируемых в различные моменты времени. Наиболее эффективными здесь представляются операции слежения, базирующиеся на свойстве пространственно-временной непрерывности движения реальных объектов.

В итоге получаем, что для когнитивной интерпретации изображений априорно неопределённых сцен, включающей обнаружение объектов, выявления их «намерений» по динамическим параметрам, их классификацию (кластеризацию) на группы с близкими свойствами и различения, может оказаться достаточным ограниченное множество P^{rcg} яркостных, геометрических, топологических и траекторных характеристик-признаков визуальных примитивов.

Безэталонная классификация

Все известные методы кластеризации и классификации, включая так называемое «обучение без учителя», используют какую-то априорную информацию о предполагаемых классах объектов. В условиях полной априорной неопределённости, для которых и предназначается когнитивный видеоавтомат, структурирование множества выявленных примитивов и объектов на классы требует специальной операции, не зависящей от априорных эталонов. Реализовать её возможно посредством статистического анализа вычисленных характеристик примитивов.

Естественная логика такой «безэталонной» классификации исходит из очевидного смысла понятия «класса» как группы объектов с похожими свойствами. Данные о таких группах в явном виде отображаются признаковыми гистограммами $\Gamma^u(F^u)$ – распределениями количества примитивов Θ по значениям вычисленных для них характеристик $F^u: \{f_{\min}^u, \dots, f_{\max}^u\}$.

В качестве указателей границ классов здесь предлагается рассматривать локальные минимумы Lm_v^u характеристических ги-

РАЗДЕЛ 2. ОБРАБОТКА СИГНАЛОВ И ДАННЫХ

стограмм. Классы объектов, выявляемые на Γ^μ , естественно называть «простыми» классами. Каждый найденный на каждой гистограмме Γ^μ интервал Δf^μ значений признака $\Delta f^\mu = \overline{f_{\min}^\mu, f_{\max}^\mu}$ является областью существования простого класса Θ_K^μ объектов, выявленных по признаку f^μ .

Сложные - «комбинаторные» - классы, описываемые несколькими признаками, могут быть выявлены как «пересечения» простых классов. Так, для двух простых классов Θ_σ^α и Θ_ζ^β (выявленных на интервалах σ и ζ признаков α и β) условием образования комбинаторного класса $\Theta_{\sigma\zeta}^{\alpha\beta}$ является $\Theta_\sigma^\alpha \cap \Theta_\zeta^\beta \neq \emptyset$, иначе - класса $\Theta_{\sigma\zeta}^{\alpha\beta}$ нет.

По выявленным таким образом классам легко могут быть вычислены объективные количественные свойства наблюдаемой заранее неизвестной сцены: число классов объектов, количество объектов в классе, матожидание и разброс значений признаков внутри классов, «расстояния» между классами и т.п.

Принципиальным преимуществом операции безэталонной классификации является возможность реализации в «интеллектуальных» информационных автоматах действительного самообучения при накоплении описаний выявляемых (классов) объектов. При этом результаты обучения по узнаванию ранее запомненных, их коррекции и/или выявлению новых формируются по «величине» рассогласования описаний, полученных на текущем и предыдущем цикле анализа.

Заключение

С целью принципиального повышения качества систем технического зрения и приближения их информационной эффективности к уровню биологического восприятия ищется обобщенная функция интерпретации априорно неопределенных сцен, преобразующая видеосигнал в данные о всех, в общем случае неизвестных, наблюдаемых объектах.

Искомое решение, названное когнитивным видеоавтоматом, строится на предположительно универсальном представлении изображений множеством визуальных примитивов - пространственно и амплитудно связанных групп пикселей. Результирующее описание сцены строится на вычисленных по изображению характеристиках примитивов. При

этом структурирование множества примитивов на «объектные» группы и классы объектов реализуется операцией безэталонной когнитивной классификации, выявляющей границы классов по локальным минимумам характеристических гистограмм.

Элементы разрабатываемой когнитивной технологии машинного зрения по обнаружению и классификации объектов различных классов, выявляемыми как группы визуальных примитивов, внедрены в медицинской диагностике [8] и использованы в комплексах видеонаблюдения [9,10].

Дальнейшее развитие когнитивного машинного зрения предполагает реализацию в СТЗ функции самообучения и построение на основе самообучающегося видеоавтомата самостоятельно перемещающегося в сложной изменяющейся среде автономного робота.

СПИСОК ЛИТЕРАТУРЫ

1. Методы компьютерной обработки изображений // под ред. В.А. Сойфера. - 2-е изд., испр. - М.: Физматлит, 2003. - 784 с.
Состояние и перспективы развития исследований в области обработки и распознавания видеоинформации (аналитический обзор) [Электронный ресурс] / Ю.И. Журавлев, К.В. Рудаков, С.И. Гуров и др.; январь, №1, 2005 - Режим доступа: <http://www.techno.edu.ru/16001/db/msg/22358>.
2. Потапов, А.С. Распознавание образов и машинное восприятие: Общий подход на основе принципа минимальной длины описания / А.С. Потапов. - СПб.: Политехника, 2007. - 548 с.
3. Визильтер, Ю.В. Обработка и анализ изображений в задачах машинного зрения: Курс лекций и практических занятий / Ю.В. Визильтер, С.Ю. Желтов, А.В. Бондаренко, М.В. Ососков, А.В. Моржин. - М.: Физматкнига, 2010. - 672 с.
4. Марр, Д. Зрение. Информационный подход к изучению представления и обработки зрительных образов // пер. с англ. - М.: Радио и связь, 1987. - 400 с.
5. Саниев, К.Б. Распознавание сигналов изображений // Радиотехника. Апрель 2007, № 4. с. 9-15.
6. Пат. 2413301 РФ, МПК G06K 9/62, G06T 1/00.
7. Саниев, К.Б. Автоматическая интерпретация сигналов изображений / К.Б. Саниев // Вестник МЭИ, №4, 2012 г., с. 52-57.
8. Автандилов, Г.Г. Плоидометрия в повышении качества патогистологической диагностики / Г.Г. Автандилов, К.Б. Саниев // Архив патологии (Arkhiv patologii), 3, том 64, май-июнь, М.: «Медицина», 2002. с.31-33.
9. Саниев, К.Б. Некоторые аспекты построения интеллектуальных комплексов видеонаблюдения / К.Б. Саниев // Научно-техн. журнал

ОПРЕДЕЛЕНИЕ КООРДИНАТ ОЧАГА ВЗРЫВА МНОГОТОЧЕЧНОЙ ОПТИКО-ЭЛЕКТРОННОЙ СИСТЕМОЙ НА ОСНОВЕ МЕТОДА ЦЕНТРА ТЯЖЕСТИ

«Цифровая Обработка Сигналов», №2, 2011 г. с. 27-31.

10. Саниев, К.Б. Предварительный модельный эксперимент по обнаружению в потоковых сигналах изображений объектов, заданных нечёткой информацией о габаритах и конфигурации / К.Б. Саниев // Сб. тр. ОАО «Концерн

радиостроения «ВЕГА», № 2(5), 2011, с.197-204.

Ведущий научный сотрудник, **Саниев К. Б.**, к.т.н., 8 (906) 703-47-91, kbsaniev@concord.ru – ЗАО «МНИТИ» (Московский научно-исследовательский телевизионный институт), 8 (499) 787-06-81, saniev@mniti.ru,

УДК: 654.924.5

ОПРЕДЕЛЕНИЕ КООРДИНАТ ОЧАГА ВЗРЫВА МНОГОТОЧЕЧНОЙ ОПТИКО-ЭЛЕКТРОННОЙ СИСТЕМОЙ НА ОСНОВЕ МЕТОДА ЦЕНТРА ТЯЖЕСТИ

С.А. Лисаков, А.Н. Павлов, Е.В. Сыпин

В статье рассматривается применение метода центра тяжести для расчета координат очага взрыва в многоточечной системе, состоящей из нескольких не координатно-чувствительных оптико-электронных датчиков, регистрирующих только поток оптического излучения. Предложенный метод апробирован методами компьютерного моделирования. Оценена относительная погрешность определения координат очага взрыва.

Ключевые слова: взрывозащита, многоточечная оптико-электронная система, координаты очага взрыва, метод центра тяжести.

Введение

В мировой практике усовершенствование оборудования взрывозащиты ведется по пути создания принципиально новых средств – автоматических систем, способных локализовать взрыв на начальной стадии развития.

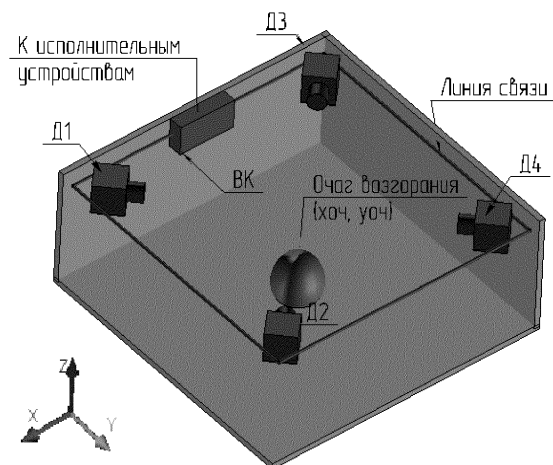
Для обеспечения максимального быстрого действия такие системы строятся на базе оптико-электронных датчиков, регистрирующих факт возникновения возгорания. Получение при этом дополнительной информации о пространственном расположении очага взрыва позволяет значительно повысить эффективность и рентабельность применения автоматической системы взрывоподавления [1 – 5].

В настоящее время на кафедре методов и средств измерений и автоматизации (МСИА) Бийского технологического института ведется разработка многоточечной оптико-электронной системы определения координат очага взрыва (МОЭС). В основе работы системы лежит многоточечный метод контроля, позволяющий обеспечить охрану объектов сложной конфигурации. Система построена на основе распределенной сети простых оптико-электронных датчиков интегрального типа, которые регистрируют только поток оптического излучения. Зная пространственное расположение датчиков и значения их выходных сигналов, можно определить

координаты источника оптического излучения внутри охраняемой зоны.

Постановка задачи

На рисунке 1 приведена схема размещения МОЭС, состоящей из четырех датчиков, в охраняемом помещении прямоугольной фор-



Д1- Д4 – оптико-электронные датчики;
VK – вычислительный комплекс;

$x_{оч}$, $y_{оч}$ – координаты очага взрыва, м.

Рисунок 1 - Схема размещения МОЭС на охраняемом объекте

мы. Для охраны объектов более сложной формы количество датчиков может быть увеличено. Сигналы с датчиков через линию