ИКОНИКА – НАУКА ОБ ИЗОБРАЖЕНИИ

УДК 621.391

ТЕОРЕТИКО-ИНФОРМАЦИОННЫЙ ПОДХОД К ВВЕДЕНИЮ ОБРАТНЫХ СВЯЗЕЙ В МНОГОУРОВНЕВЫХ СИСТЕМАХ МАШИННОГО ЗРЕНИЯ

© **2007 г.** **А. С. Потапов**, канд. техн. наук

НПК “Государственный оптический институт им. С.И. Вавилова”, Санкт-Петербург

Рассмотрена проблема введения обратных связей в многоуровневых системах машинного зрения. На основе теоретико-информационного анализа показано, что необходимость обратных связей в таких системах обусловлена тем, что на разных уровнях оптимизируются отдельные компоненты общего для всей системы критерия качества. При этом принятие решений на ранних этапах анализа осуществляется при использовании только части компонентов этого критерия, что приводит к неоптимальности этих решений. Предложен подход к введению обратных связей как средства достижения глобального оптимума информационного критерия качества путем итеративного улучшения решения. На основе полученных результатов обосновывается недостаточность информации, содержащейся в контурах, для робастного построения структурных элементов.

Коды OCIS: 150.0150, 3000.32960.

Поступила в редакцию 04.04.2007.

1. Введение

Важность концепции обратной связи хорошо известна. Наличие обратной связи в направлении от среды в биологических системах во многом определяет их способность к выживанию. Адаптивные и обучающиеся робототехнические системы также создаются на основе обратной связи, возникающей за счет подсистем машинного восприятия. Однако обратная связь может быть установлена не только от среды к кибернетической системе, но также и между подсистемами самой системы, где она играет не меньшую роль.

В статье рассматриваются системы машинного зрения, которые в настоящее время в большинстве случаев разрабатываются как многоуровневые [1]. В многоуровневых системах сигнал поступает на вход – нижний уровень, и далее распространяется на более высокие уровни, подвергаясь все более сложному анализу. Однако известно [2], что в биологических системах восприятия существуют также многочисленные обратные связи в направлении от верхних уровней к нижним. При этом верхние уровни за счет вовлечения более широкой контекстной информации управляют процессом анализа, осуществляемого на нижних уровнях.

В области машинного зрения существуют методы, в которых поток информации вводится в направлении от верхних уровней к нижним. Преимущественно эти методы относятся к методам, основанным на знаниях, и обычно применяются в случаях, когда содержание входного сигнала на верхнем уровне достаточно хорошо известно априори и его требуется лишь уточнить (см., например, [3]).

Когда речь идет об обратной связи от среды, смысл этой связи достаточно ясен. Например, в системах автоматического управления с обратной связью обычно решается задача поддержания заданного значения регулируемой величины, описывающей состояние управляемого объекта (среды) на основе информации, содержащейся во входном сигнале [4, стр. 26]. Информация, поступающая по обратной связи, порождается самой средой.

Однако при введении обратных связей в многоуровневые системы машинного зрения оказывается необходимым определить не только реакцию на сигналы, поступающие по этим связям, но и сам характер связей, т. е. определить, какую именно информацию следует передавать “вниз” и с какой целью. Естественно, вводимые обратные связи должны облегчать решение самой задачи восприятия. Эту задачу мы трактуем как задачу индуктивного вывода и проводим ее исследование в рамках теоретико-информационного подхода [5]. В данной работе показано, как теоретико-информационный подход может быть использован для объяснения назначения обратных связей в многоуровневых системах восприятия и разработки методики их использования.

2. Существующие подходы

2.1. Теория адаптивного резонанса

Одно из первых детальных исследований роли обратных связей в восприятии привело в конце 1970-х гг. к разработке теории адаптивного резонанса. Сама концепция адаптивного резонанса нагляд-

нее всего выражается в нейронной сети типа ART[1], предложенной С. Гроссбергом [6, 7].

Данная сеть состоит из нескольких (в простейшем случае из двух) слоев. Первый слой является входным. Каждый его нейрон представляет некоторый признак объекта. Во втором слое нейроны соответствуют классам объектов. Каждый нейрон второго слоя получает информацию от нейронов первого слоя и активируется тогда, когда признаки соответствуют классу объектов, связанному с данным нейроном. Это похоже на сеть прямого распространения, однако от нейронов второго слоя к первому слою также идут связи. Если какой-то нейрон второго слоя активен, то он осуществляет поощрение нейронов первого слоя, паттерн активности которых соответствует признакам эталонного объекта данного класса. Таким образом, снизу вверх передается текущая информация, а сверху вниз передается ожидание.

Сеть ART является рекуррентной, т. е. изменение активности нейронов первого слоя сказывается на активности нейронов второго слоя, которая, в свою очередь, влияет на активность нейронов первого слоя. Если какие-то нейроны первого слоя не получают подкрепления со второго слоя, то их активность постепенно затухает. И наоборот, активность нейронов, получивших исходно слабый сенсорный сигнал, может быть усилена, если сверху к ним поступает сильное подкрепление (начальная активность нейронов второго слоя соответствует априорным ожиданиям).

В такой сети сохраняются лишь резонансные состояния, т. е. такие паттерны активности нейронов на двух уровнях, которые поддерживают друг друга. Таким образом, результат восприятия складывается из ожиданий и новой информации. При этом осуществляется коррекция активности нейронов нижнего уровня, что позволяет восстанавливать искаженную и пропущенную информацию.

Время на установление резонанса может служить объяснением, почему восприятие оказывается более медленным процессом, чем прямое прохождение сигналов от органов чувств к мозгу.

Существует ряд иллюзий восприятия, которые объясняются этой теорией [2]. Подкрепление, идущее с верхних уровней нейронной сети ART, позволяет также моделировать управление вниманием: подкрепляются признаки тех объектов в перцептивном поле, на которые направлено внимание. Однако самое важное заключается в том, что эта модель раскрыла механизмы восприятия в условиях зашумленности и неопределенности сенсорных сигналов, что является весьма актуальным для систем машинного восприятия.

[1] ART – Adaptive resonance theory (теория адаптивного резонанса).

К сожалению, концепция адаптивного резонанса получила развитие преимущественно в рамках нейросетевого подхода [7], хотя ее значение существенно шире. Сама нейронная сеть ART (как и ее усовершенствования) при практическом использовании обладает некоторыми недостатками. Архитектура подобной сети удобна для распознавания единичного объекта с фиксированным числом признаков. Однако эта архитектура не подходит для анализа сложной зрительной сцены или речевых высказываний неопределенной длины.

Сама архитектура сети с обратными связями не говорит о том, как следует эти связи настраивать, чтобы добиться оптимального функционирования. Хотя сеть ART позволяет прояснить ряд вопросов в восприятии и реализует ряд фундаментальных концепций, построение оптимальной технической системы восприятия выходит за рамки нейросетевой парадигмы.

2.2. Байесовский подход

Рассмотрим проблему введения обратных связей в рамках байесовского подхода, который хорошо зарекомендовал себя и широко используется при интерпретации изображений [8].

Пусть D – исходные данные, а H – некоторая гипотеза об их содержании. Тогда в рамках байесовского подхода лучшей гипотезой будет та, для которой максимизируется апостериорная вероятность $P(H|D)$, выражающаяся, согласно правилу Байеса, через правдоподобие данных $P(D|H)$ и априорную вероятность гипотезы $P(H)$:

$$P(H|D) = \frac{P(D|H)P(H)}{P(D)}, \qquad (1)$$

где априорная вероятность данных $P(D)$ не зависит от гипотезы и в процессе выбора может не учитываться.

Мы рассматриваем многоуровневые системы вывода, поэтому следует записать

$$\begin{gathered} P(H_1|D) = P(D|H_1)P(H_1)/P(D), \\ P(H_2|H_1) = P(H_1|H_2)P(H_2)/P(H_1), \\ \cdots \\ P(H_N|H_{N-1}) = P(H_{N-1}|H_N)P(H_N)/P(H_{N-1}). \end{gathered} \qquad (2)$$

Здесь H_i – это гипотеза i-го уровня, для выбора которой используется в качестве исходных данных гипотеза $(i-1)$-го уровня.

В случае многоуровневой системы следует максимизировать совместную апостериорную вероятность $P(H_1, H_2, \ldots, H_N|D)$, где $(H_1, H_2, \ldots, H_N)$ – общая гипотеза о содержании сенсорного сигнала, которая представлена в виде совокупности частных гипотез разных уровней.

Рассмотрим два уровня. Тогда $P(H_1, H_2|D) = P(H_2|H_1)P(H_1|D)$. Предположим теперь, что на первом уровне выбирается некоторая лучшая гипотеза H_1^*, для которой достигается максимум $P(H_1|D)$. Тогда далее можно максимизировать величину

$$P(H_1, H_2|D) = P(H_2|H_1^*)P(H_1^*|D), \qquad (3)$$

где $P(H_1^*|D) = \text{const}$.

Теперь отметим следующее: максимизация $P(H_1|D)$ необязательно обеспечивает максимизацию $P(H_1, H_2|D)$, так как при выбранном H_1^* другой множитель ($P(H_2|H_1^*)$) может быть небольшим.

Однако перебор всех возможных пар (H_1, H_2) для поиска максимума $P(H_1, H_2|D)$ является неэффективным с вычислительной точки зрения. При увеличении количества исходных данных (например, увеличении размеров изображения) пространство гипотез экспоненциально растет и возникает проблема комбинаторного взрыва. Таким образом, с одной стороны, нужно избежать перебора всех возможных комбинаций гипотез, а с другой – обеспечить поиск максимума.

С этой целью, как правило, с нижних уровней вверх подают не одну лучшую гипотезу, а несколько альтернатив. Останется та гипотеза H_1^*, которая не просто дает максимум $P(H_1|D)$, но и для которой найдется такая гипотеза H_2^*, что множитель $P(H_2^*|H_1^*)$ будет достаточно большим (таким, что $P(H_1^*, H_2^*|D)$ будет больше, чем у альтернатив). Таких альтернатив, однако, не может быть рассмотрено слишком много, так как для каждой гипотезы нижнего уровня нужно будет строить гипотезы следующего уровня. Следовательно, выбираются наиболее различающиеся альтернативы. Из близких по содержанию гипотез выбирается лучшая, но эта гипотеза также может оказаться неоптимальной на следующем уровне.

В связи с этим можно ввести следующую операцию, которая, правда, в байесовском подходе обычно не реализуется. После того как для некоторой гипотезы H_1^* найдена лучшая гипотеза H_2^*, можно найти такую гипотезу H_1', которая максимизирует второй множитель – $P(H_2^*|H_1)$, после чего приблизить H_1^* к H_1'. Условно можно записать $H_1^* = (1 - \varepsilon)H_1^* + \varepsilon H_1'$, где ε – некоторый множитель. Ожидаемая гипотеза H_2^* говорит, какова должна быть гипотеза предыдущего уровня. Итеративная взаимная корректировка H_1 и H_2 с целью максимизации произведения и дает оптимальные значения H_1 и H_2.

Таким образом, в байесовском подходе, как и в подходе на основе адаптивного резонанса, может активироваться по несколько гипотез на каждом уровне, а самому эффекту адаптивного резонанса соответствует итеративная оптимизация произведения $P(H_2|H_1)P(H_1|D)$ за счет обоих сомножителей, что заменяет исчерпывающий поиск в пространстве гипотез.

Байесовский подход является математически корректным, и с его помощью можно достаточно четко проследить причины возникновения и смысл обратных связей, однако его применение может быть затруднено из-за необходимости вычисления вероятностей для гипотез, обладающих сложной структурой. Например, оценивание $P(H_1|D)$, где D – изображение, а H_1 – совокупность выделенных на нем контуров, представляется крайне проблематичным. Применение байесовского подхода в многоуровневых системах не вызывает трудностей только начиная со средних уровней (фонем или букв) при распознавании устной или письменной речи или при распознавании узкого класса объектов по векторам признаков фиксированного размера.

Также нетривиальным может оказаться задание априорных вероятностей, которые и определяют ожидания. Если нет абсолютно четкого ожидания содержания входного сигнала, то задание априорных вероятностей, особенно на нижних уровнях системы восприятия, оказывается проблематичным. Например, как задать априорные вероятности для тех или иных контуров на изображении, когда нет ожидания конкретного контура в конкретном месте? Проблема априорных вероятностей является очень глубокой [9] и здесь не рассматривается.

3. Теоретико-информационный подход к введению обратных связей

Мы предлагаем решение проблемы обратных связей в рамках теоретико-информационного подхода, основанного на принципе минимальной длины описания. В словесной формулировке этот принцип гласит [9]: среди всех моделей, объясняющих данные, следует выбрать ту, которая минимизирует сумму длины описания данных посредством модели и длины описания самой модели. Длины описания, как правило, вычисляются как длины строк, кодирующих данные посредством модели и саму модель. Схема кодирования определяется представлением информации, соответствующим метамодели данной предметной области.

Таким образом, выбор гипотезы осуществляется путем минимизации длины описания

$$L(D, H) = L(D|H) + L(H), \qquad (4)$$

что есть информационный аналог правила Байеса. Однако в отличие от байесовского подхода теоретико-информационный подход позволяет оценивать длину описания гипотез со сложной структурой.

Рассмотрим теоретико-информационный подход к введению обратных связей в системах интерпре-

тации изображений. Сам вопрос о возможных представлениях информации в этих системах разбирать не будем, а используем представления, описанные в [5].

Будем считать, что изображение $f(x, y)$, заданное в области G, т. е. f: $G \to R$, описывается как совокупность непересекающихся областей $G = \bigcup_i G_i$. Содержимое каждой области (сужение изображения f на область G_i, т. е. $f|_{G_i}$) описывается индивидуальной стохастической моделью. В простейшем случае интенсивности пикселов в области G_i описываются как статистически независимые и одинаково распределенные (функция распределения вероятностей разная для разных областей). В более сложном случае могут учитываться пространственные зависимости интенсивностей пикселов. Подобное представление обосновано тем, что наблюдаемые сцены состоят из видимых поверхностей, каждая из которых обладает индивидуальной картой отражательной способности.

Помимо интенсивностей пикселов внутри каждой области, также необходимо описать положение этих областей. Это можно сделать через описание их границ ∂G_i.

Тогда длина описания изображения в рамках некоторой гипотезы о разбиении этого изображения на области представляется в виде суммы

$$DL = \sum_i \left(L\left(f|_{G_i} \right) + L\left(\partial G_i \right) \right). \qquad (5)$$

Далее возникает вопрос о кодировании контуров (границ областей). В рассматриваемом нами представлении контуры описываются через совокупность структурных элементов, каждый из которых соответствует некоторому сегменту контура. В описании необходимо указывать границы сегментов (что соответствует углам и пересечениям на контуре) и форму линии, описывающей сам сегмент (например, параметры прямой или дуги окружности или эллипса). Поскольку сам контур описывается структурными элементами с некоторыми ошибками, то эти ошибки также нужно закодировать, чтобы вычислить корректную длину описания.

Пусть $s_j^{(i)}$ – j-й структурный элемент на i-м контуре ∂G_i. Тогда

$$L\left(\partial G_i \right) = \sum_j \left(L\left(\partial G_i|_{s_j^{(i)}} \right) + L\left(s_j^{(i)} \right) \right). \qquad (6)$$

Далее описание структурных элементов $s_j^{(i)}$ может вестись путем их группирования по сходству и регулярности пространственного положения [5, стр. 316], но мы ограничимся первыми тремя уровнями.

Таким образом, для вычисления точного значения критерия качества (5) для каждой гипотезы о разбиении изображения на области необходимо вычислить значение (6), для чего следует осуществить построение структурных элементов на основе границы этих областей (т. е. найти такие $s_j^{(i)}$, для которых достигается минимум $L(\partial G_i)$ для каждой совокупности областей G_i). Это эквивалентно перебору всех возможных пар гипотез H_1, H_2 в байесовском подходе.

Как уже говорилось, это неэффективно с вычислительной точки зрения. С другой стороны, минимизация слагаемого $L\left(\partial G_i|_{s_j^{(i)}} \right)$ без учета слагаемого $L(s_j^{(i)})$ ведет к построению большого количества областей с сильно изрезанными границами, поэтому следует использовать некую простую оценку длины описания $L(\partial G_i)$, для чего не требуется построение гипотез следующего уровня. Иначе говоря, вместо изощренного кодирования контуров через структурные элементы можно применить несколько менее эффективное, но гораздо более быстрое кодирование контуров, например, цепное кодирование, при котором указываются первая точка контура и направление от каждой предыдущей точки к последующей.

Таким образом, для построения областей может использоваться критерий качества

$$DL_0 = \sum_i \left(L\left(f|_{G_i} \right) + L_0\left(\partial G_i \right) \right), \qquad (7)$$

где $L_0(\partial G_i)$ – экспресс-оценка значения $L(\partial G_i)$, причем $L_0(\partial G_i) \geq L(\partial G_i)$.

Далее при фиксированных контурах может осуществляться построение структурных элементов с минимизацией критерия качества (6).

Очевидно, найденная путем минимизации критерия (7) гипотеза об областях на изображении может не приводить к минимизации общего критерия (5). Вопрос о проведении границы между какими-то двумя областями может быть неоднозначным, когда длина описания на этом уровне почти не меняется. Такие случаи могут быть достаточно частыми, особенно на зашумленных изображениях. Разрешение этой неоднозначности может быть осуществлено на следующем уровне, если принять во внимание длину описания границ областей. Таким образом, при данном подходе также следует передавать на следующие уровни несколько альтернативных гипотез.

На следующем уровне описание контуров посредством структурных элементов неточно, т. е. найденные контуры отклоняются от идеальных отрезков и сегментов. Здесь возможны два способа описания этих отклонений. Во-первых, могут опи-

сываться сами отклонения, т. е. описывается отклонение каждой точки контура от точек, задаваемых структурными элементами, что выражается через слагаемое $L\left(\partial G_i\big|_{s_j^{(i)}}\right)$. Во-вторых, сам контур может быть скорректирован. При этом меняются границы областей, а значит, меняется и слагаемое $L(f|_{G_i})$. Если из-за плавного перехода или больших шумов граница области определяется неточно, то смещение этой границы не оказывает большого влияния на длину описания содержимого областей, однако может сильно влиять на длину описания границ.

Собственно обратным связям соответствует корректировка границ областей. Структурные элементы задают ожидания положения контуров, и контур может притягиваться к этому ожиданию при условии, что суммарная длина описания уменьшается, т. е. увеличение длины описания содержимого области компенсируется уменьшением длины описания контуров.

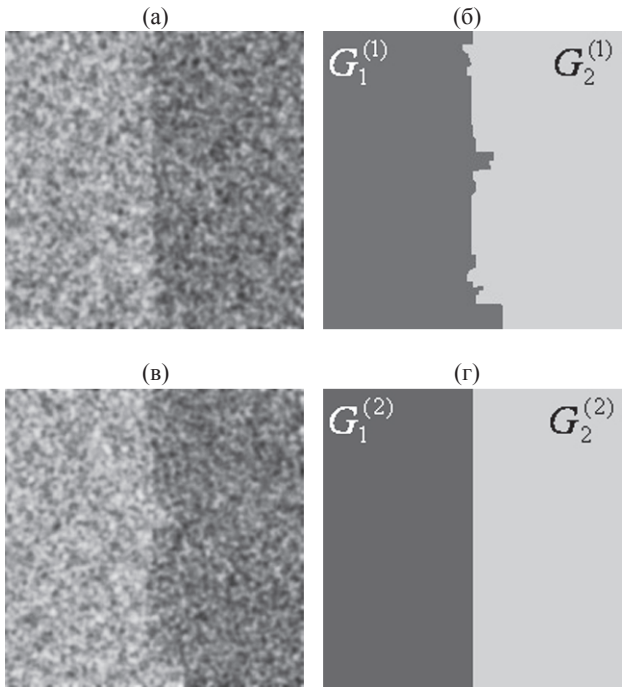


Рис. 1. Разделение изображения на области с использованием обратных связей. а – синтезированное изображение, состоящее из двух областей, разделенных прямолинейной (вертикальной) границей; б – области, найденные путем минимизации критерия качества (7), в – изображение, синтезированное с использованием областей (б), г – области, найденные для изображения (а) в результате коррекции областей (б) с помощью обратных связей.

Экспериментальная проверка

Проверим, действительно ли одновременный учет обоих слагаемых $L(f|_{G_i})$ и $L(\partial G_i)$ в уравнении (5), реализующийся через введение обратных связей от уровня структурных элементов к уровню контуров, позволяет получить адекватный результат. В качестве оценки $L_0(\partial G_i)$ будем использовать $2\|\partial G_i\|$, где $\|\partial G_i\|$ – число точек в контуре (для определения положения каждой точки на контуре требуется знать направление от предыдущей точки, которое можно закодировать двумя битами для четырехсвязного контура).

На рис. 1((а) и (в)) представлены два синтезированных изображения, для каждого из которых в результате оптимизации критерия (7) находятся области, приведенные на рис. 1б. Для изображения (а) правильным было бы разделение, приведенное на рис. 1г, однако из-за локальных флуктуаций яркости найденная граница смещена относительно истинной границы. Обозначим изображения (а) и (в) через f_1 и f_2 соответственно. В табл. 1 и 2 приведены длины описаний изображений f_1 и f_2 в зависимости от выделенных областей.

Разница между оценками длин описания границ в двух случаях оказывается равной примерно 72 би-

Таблица 1. Длины описания содержимого изображения f_1, суженного на области $G_1^{(1)}$, $G_2^{(1)}$, $G_1^{(2)}$, $G_2^{(2)}$

Вариант разделения на области	G_1		G_2		$L\left(f_1\big\|_{G_1}\right)+L\left(f_1\big\|_{G_2}\right)$
	Энтропия	$L\left(f_1\big\|_{G_1}\right)$, бит	Энтропия	$L\left(f_1\big\|_{G_2}\right)$, бит	
$G^{(1)}$	6,362	129826	6,285	123140	252966
$G^{(2)}$	6,361	127219	6,295	125906	253125

Таблица 2. Длины описания содержимого изображения f_2, суженного на области $G_1^{(1)}$, $G_2^{(1)}$, $G_1^{(2)}$, $G_2^{(2)}$

Вариант разделения на области	G_1		G_2		$L\left(f_2\big\|_{G_1}\right)+L\left(f_2\big\|_{G_2}\right)$
	Энтропия	$L\left(f_2\big\|_{G_1}\right)$, бит	Энтропия	$L\left(f_2\big\|_{G_2}\right)$, бит	
$G^{(1)}$	6,328	129142	6,236	122178	251320
$G^{(2)}$	6,343	126859	6,313	126242	253101

там. В то же время при получении более точных оценок в результате построения структурных элементов разница длин описания $L(\partial G^{(1)}) - L(\partial G^{(2)})$ возрастает до 232 бит. Пользуясь данными, приведенными в таблицах, получаем

$$\left[L\left(f_1\big|_{G_1^{(1)}}\right)+L\left(f_1\big|_{G_2^{(1)}}\right)+L_0\left(\partial G^{(1)}\right)\right]-$$
$$-\left[L\left(f_1\big|_{G_1^{(2)}}\right)+L\left(f_1\big|_{G_2^{(2)}}\right)+L_0\left(\partial G^{(2)}\right)\right]=-97 \text{ бит},$$
$$\left[L\left(f_2\big|_{G_1^{(1)}}\right)+L\left(f_2\big|_{G_2^{(1)}}\right)+L_0\left(\partial G^{(1)}\right)\right]-$$
$$-\left[L\left(f_2\big|_{G_1^{(2)}}\right)+L\left(f_2\big|_{G_2^{(2)}}\right)+L_0\left(\partial G^{(2)}\right)\right]=-1709 \text{ бит},$$
$$\left[L\left(f_1\big|_{G_1^{(1)}}\right)+L\left(f_1\big|_{G_2^{(1)}}\right)+L\left(\partial G^{(1)}\right)\right]-$$
$$-\left[L\left(f_1\big|_{G_1^{(2)}}\right)+L\left(f_1\big|_{G_2^{(2)}}\right)+L\left(\partial G^{(2)}\right)\right]=63 \text{ бит},$$
$$\left[L\left(f_2\big|_{G_1^{(1)}}\right)+L\left(f_2\big|_{G_2^{(1)}}\right)+L\left(\partial G^{(1)}\right)\right]-$$
$$-\left[L\left(f_2\big|_{G_1^{(2)}}\right)+L\left(f_2\big|_{G_2^{(2)}}\right)+L\left(\partial G^{(2)}\right)\right]=-1549 \text{ бит}.$$

Таким образом, более точная оценка длины описания границ, получаемая в результате построения структурных элементов, позволяет выбрать разделение (г) для f_1, сохранив выбор разделения (б) для f_2.

Здесь стоит отметить, что при разделении изображения на области структурные элементы еще не построены, поэтому в качестве начальной гипотезы строится в обоих случаях разделение, приведенное на рис. 1б. Обычно полученные контуры используются для дальнейшего анализа без обращения к исходным интенсивностям пикселов. Однако в нашем примере для обоих изображений f_1 и f_2 контуры были построены одинаковыми, т. е. на контурном уровне эти изображения далее интерпретировались бы как эквивалентные. Преимущество обратных связей видно также и при работе с реальными изображениями (см. рис. 2).

Суть обратных связей заключается в том, что при последующем построении структурных элементов выдвигаются новые гипотезы о положении контуров и для определения качества этих гипотез необходимо возвращаться на предыдущий уровень для получения оценок $L(f|_{G_i})$. Иными словами, контуры не содержат всей информации, необходимой для робастного построения структурных элементов. Этот вывод можно распространить и на следующие уровни зрительного восприятия. Более того, с увеличением уровней в иерархических системах при отсутствии обратных связей неоптимальность решений, принимаемых на нижних уровнях, усиливается на последующих уровнях, что может приводить к катастрофическому падению эффективности всей системы в целом.

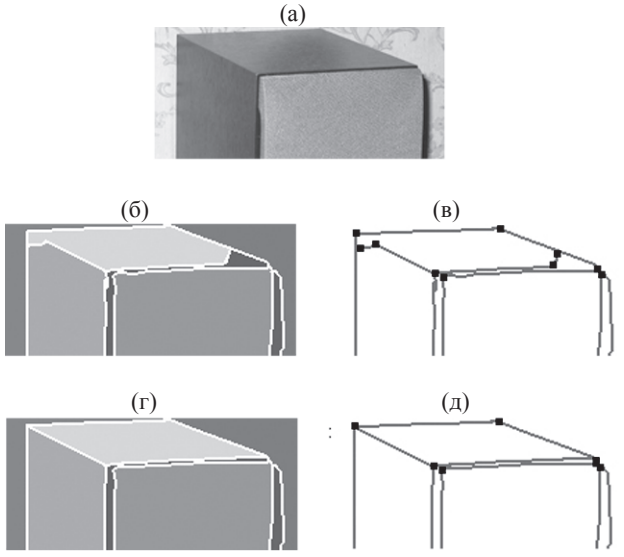


Рис. 2. Выделение областей и структурных элементов на изображении. а – фрагмент исходного изображения, б – результат выделения областей путем минимизации целевой функции (7), в – результат построения структурных элементов на основе границ областей (б), г – результат улучшения областей (б) за счет обратных связей от уровня структурных элементов, д – структурные элементы, построенные в результате итеративной оптимизации критерия качества (5) вместе с областями (г).

Заключение

В работе рассмотрен теоретико-информационный подход к введению обратных связей в системах машинного зрения. Этот подход является более строгим, чем подход на основе теории адаптивного резонанса, и позволяет определить оптимальную форму обратных связей при заданном представлении информации. На практике этот подход также предпочтительнее, чем байесовский, поскольку не требует явного задания априорных вероятностей гипотез (вид которых, как правило, не известен) и позволяет привлекать более сложные описания изображений, использование которых в рамках байесовского подхода затруднительно из-за невозможности задания соответствующих стохастических моделей.

Проведенное исследование показывает, что данный подход позволяет повысить робастность методов построения иерархических структурных описаний изображений, причем проблема введения обратных связей становится центральной при увеличении степени априорной неопределенности содержания изображений.

Для дальнейшего развития данного подхода необходимо разрабатывать конкретные механизмы итеративной оптимизации глобального критерия качества, связывающего отдельные уровни системы зрения.

ЛИТЕРАТУРА

1. *Rares A., Reinders M.J.T., Hendriks E.A.* Image Interpretation Systems // Technical Report (MCCWS 2.1.1.3.C). TU Delft. 1999. 32 p.
2. *Gove A., Grossberg S., Mingolla E.* Brightness perception, illusory contours, and corticogeniculate feedback // Visual Neuroscience. 1995. V. 12. P. 1027–1052.
3. *Liedtke C.-E., Grau O., Growe S.* Use of explicit knowledge for the reconstruction of 3-D object geometry // Int. Conf. on Computer Analysis of Images and Patterns. 1995. P. 580–587.
4. *Козлов Ю.М.* Адаптация и обучение в робототехнике. М.: Наука. Гл. ред. физ.-мат. лит., 1990. 248 с.
5. *Потапов А.С.* Распознавание образов и машинное восприятие. Общий подход на основе принципа минимальной длины описания: СПб., Политехника, 2007. 548 с.
6. *Grossberg S.* Adaptive pattern classification and universal recoding. I: Parallel development and coding of neural feature detectors & II: Feedback, expectation, olfaction, and illusions // Biological Cybernetics. 1976. V. 23. P. 121–134, 187–202.
7. *Grossberg S.* Adaptive resonance theory // Technical Report CAS/CNS TR-2000-024. In The encyclopedia of cognitive science. London: Macmillan Reference Ltd., 2003.
8. *Kopparapu S.K., Desai U.B.* Bayesian Approach to Image Interpretation // The Springer International Series in Engineering and Computer Science. 2001. V. 616. 127 p.
9. *Li M., Vitanyi P.* Philosophical Issues in Kolmogorov complexity // Proc. ICALP'92 (invited lecture). 1992. P. 1–15.