

УДК 519.6

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ АВТОСТЕРЕОСКОПИЧЕСКИХ УСТРОЙСТВ

А. А. Лукьяница
(МГУ им. М. В. Ломоносова)

Предлагаются новые типы автостереоскопических устройств — дисплея и проекционной системы. Для описания их функционирования построены математические модели, а также сконструированы эффективные алгоритмы расчета изображений, позволяющие формировать на указанных устройствах объемное световое поле. Как дисплей, так и проекционная система могут использоваться для визуализации трехмерных данных в научных исследованиях.

Введение

В настоящее время рынок устройств, отображающих визуальную информацию, близок к насыщению. Действительно, если, например, рассмотреть современные жидкокристаллические (ЖК) дисплеи, выпускаемые различными компаниями, то можно заметить, что все они имеют похожие характеристики: яркость, контрастность, углы обзора и т. п. Из-за этого при покупке дисплея в большей степени руководствуются дизайном либо именем производителя, а не техническими параметрами. Принципиально новым шагом, позволяющим достичь нового качества, может быть создание устройств, позволяющих воспроизводить трехмерную информацию.

В настоящей работе предлагаются два типа таких устройств: дисплей и проекционная система, относящиеся к разряду так называемых автостереоскопических устройств. Это означает, что для их использования не нужны дополнительные приспособления типа очков — трехмерное изображение видно естественным образом. Основным отличием предлагаемых устройств от ранее известных является то, что для формирования трехмерного светового поля используется не только специальная оптическая система, но еще и рассчитанные особым образом изображения, подаваемые на вход этой системы.

Для того чтобы наблюдатель видел объемное изображение, нужно сделать так, чтобы в левый и правый глаз попадали разные изображения. Обычно для этой цели используются специальные очки: анаглифические — с красным и синим

стеклами, поляризационные — в которых стекла для левого и правого глаза имеют ортогональную поляризацию, и т. п.

В предложенной автором модели объемного дисплея [1–3] разделение световых потоков достигается следующим образом. Дисплей состоит из двух ЖК панелей, между которыми помещен диффузор (рис. 1). На панели выводятся специальным образом рассчитанные изображения, которые являются масками для источника света. А поскольку ЖК панели разнесены по глубине на определенное расстояние, то изображения, попадающие в левый и правый глаз, пройдут через разные пиксели маски и, следовательно, будут разными.

Диффузор выполняет двойную функцию. Во-первых, он нужен, чтобы устранить муар, который возникает из-за того, что ЖК панели являются периодической структурой — пиксели разделены непрозрачной черной маской. Вторая функция диффузора — увеличить угол обзора: при изменении положения наблюдателя происходит наложение других пикселей маски, в результате чего изображение изменяется. Диффузор производит небольшое *размытие* задней маски, поэтому незначительные отклонения наблюдателя от выбранного положения остаются практически незаметными.

В предлагаемой проекционной системе [4] объемный эффект достигается за счет использования световозвращающего экрана — ретрорефлектора. Объемное изображение в этой системе создается следующим образом.

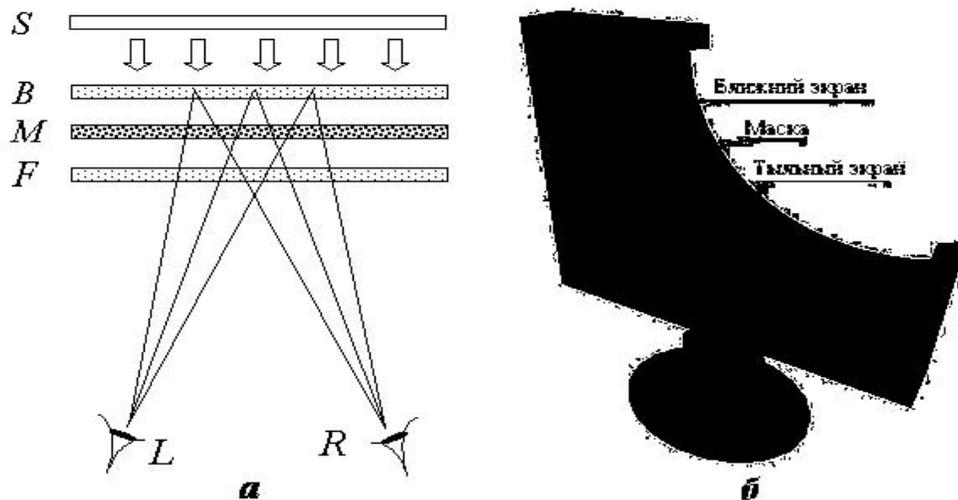


Рис. 1. Устройство автостереоскопического дисплея: *a* — вид сверху; *б* — дисплей в разрезе; *S* — тыльный осветитель; *F* и *B* — передняя и тыльная ЖК панели; *M* — маска (диффузор); *L* и *R* — левый и правый глаза наблюдателя

Рассмотрим для простоты случай, когда используются только два проектора P_L и P_R и экран S (рис. 2).

Ретрорефлектор обладает свойством отражать свет строго в направлении источника. То, что реальный ретрорефлектор отражает свет в некоторую область с экспоненциальным затуханием на периферии, можно использовать для создания объемного светового поля в достаточно широкой зоне. Анализ различных световозвращающих материалов показал, что для необходимых целей лучше всего подходит материал OLF2301, изготавливаемый фирмой 3М. Этот материал является тонкой прозрачной поликарбонатной пленкой, тыльная сторона которой состоит из вертикальных призм, как это видно из рис. 3.

Благодаря такой структуре экран обладает световозвращающими свойствами в горизонтальном направлении и является практически диффузным по вертикали. Это позволяет размещать проекторы на произвольной высоте, не обязательно на уровне глаз наблюдателя, что требуется при использовании стандартных ретрорефлекторных материалов. Закон рассеяния таким экраном хорошо моделируется нормальным распределением

$$K(\alpha) = P \exp\left(-\frac{\alpha^2}{\sigma^2}\right), \quad (1)$$

где P — величина поглощения света экраном; α — угол между источником света и наблюдателем; σ — параметр, характеризующий рас-

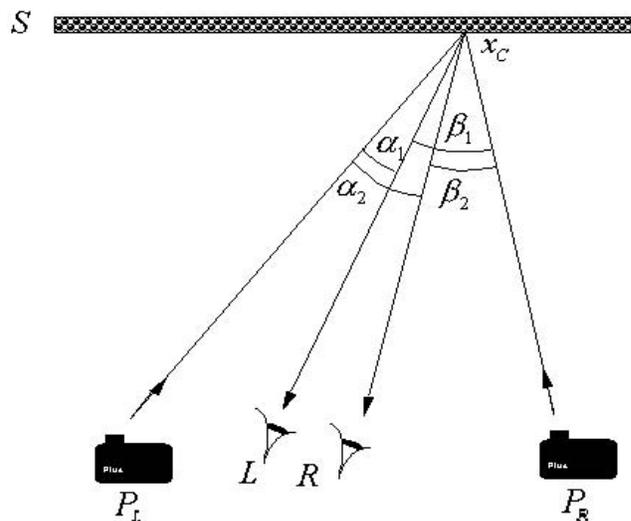


Рис. 2. Устройство проекционной системы (вид сверху)

сеивающие свойства экрана. Если на проекторы подать специальным образом рассчитанные изображения, учитывающие закон отражения (1), то наблюдатель увидит объемную картину, поскольку в разные глаза попадут изображения, отраженные под разными углами.

Таким образом, видно, что как для объемного дисплея, так и для проекционной системы нужно уметь рассчитывать подаваемые на них изображения, используя для этого модели функционирования данных устройств. Ниже приводятся математические модели работы дисплея и проекционной системы, а также эффективные алгоритмы вычисления изображений, основанные на

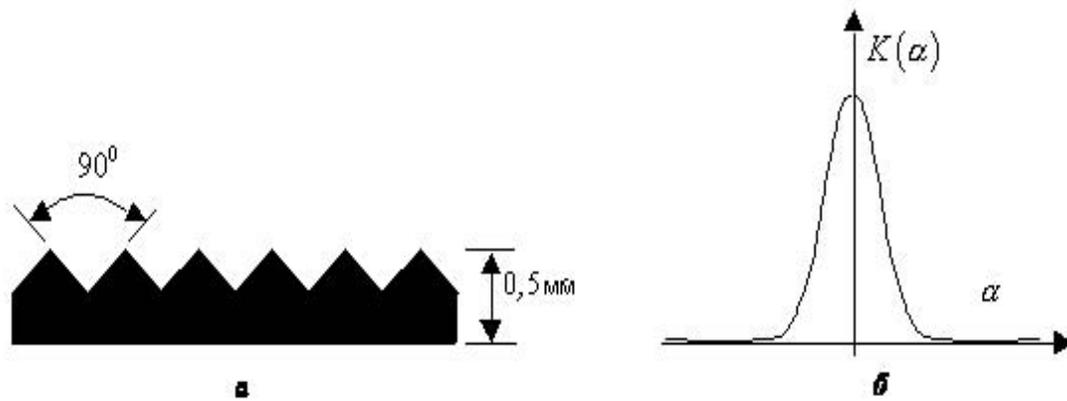


Рис. 3. Характеристики материала, используемого в качестве экрана: *a* — структура материала в горизонтальном сечении; *b* — график функции рассеяния света в зависимости от угла α между источником света и наблюдателем

искусственной нейронной сети. Достоинством предложенных алгоритмов является то, что они могут быть легко реализованы на специализированном процессоре. Использование в качестве такого процессора графического акселератора стандартной видеокарты позволило добиться ускорения более чем в 1 000 раз по сравнению с Pentium IV/3 MHz [5].

Математическая модель стереодисплея

Введем ортогональную систему координат, связанную с наблюдателем. Направим ось Ox вправо, ось Oy — вниз, а ось Oz — так, чтобы она дополняла оси Ox и Oy до правой "тройки". Пусть начало координат находится в точке, расположенной посередине между глазами наблюдателя. В выбранной системе координат левый глаз будет находиться в точке $(x_L, 0, 0)$, а правый глаз — в точке $(x_R, 0, 0)$, причем $x_L = -x_R$. Пусть экраны расположены перпендикулярно оси Oz , причем передний экран расположен на расстоянии D от начала координат, а зазор между экранами равен Δ (рис. 4).

Проведем прямую через точку, совпадающую с положением левого глаза, и некоторую точку (x, y, D) на переднем экране. Эта прямая в выбранных координатах описывается уравнением

$$\frac{X - x_L}{x} = \frac{Y}{y} = \frac{D + \Delta}{D}. \quad (2)$$

Прямая (2) пересечет задний экран в точке $(ax - \Delta x_L, ay, D + \Delta)$, где $a = (D + \Delta)/D$. В результате размытия диффузором изображе-

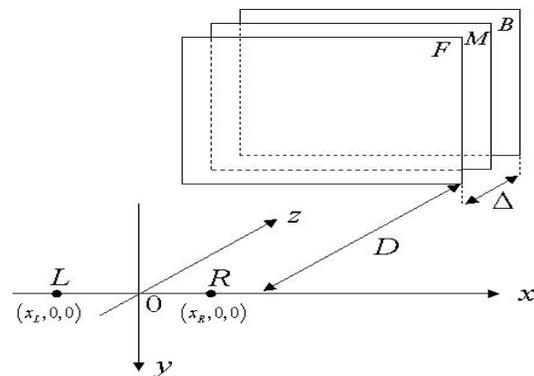


Рис. 4. Выбор системы координат

ния от тыльного экрана в левый глаз попадет изображение

$$L(x, y) = F(x, y) \times \iint_{\Omega} K(ax - \Delta x_L - \xi, ay - \eta) B(\xi, \eta) d\xi d\eta, \quad (3)$$

где $F(x, y)$ и $B(x, y)$ — интенсивности изображений соответственно на ближнем и тыльном экранах в точках (x, y, D) и $(x, y, D + \Delta)$, а

$$K(t_1, t_2) = pe^{-(t_1^2 + t_2^2)/\sigma^2} \quad (4)$$

— ядро, моделирующее действие диффузора. Здесь p — коэффициент поглощения, σ — параметр, характеризующий протяженность области рассеяния. Интегрирование в (3) проводится по области Ω , занимаемой изображением.

Проведем теперь прямую, проходящую через точку $(x_R, 0, 0)$, совпадающую с положением правого глаза, и точку (x, y, D) на переднем экране:

$$\frac{X - x_R}{x} = \frac{Y}{y} = \frac{D + \Delta}{D}. \quad (5)$$

Проведя аналогичные рассуждения, получим, что в правый глаз попадет следующее изображение:

$$R(x, y) = F(x, y) \times \int_{\Omega} K(ax - \Delta x_R - \xi, ay - \eta) B(\xi, \eta) d\xi d\eta. \quad (6)$$

Таким образом, для того чтобы найти изображения $F(x, y)$ и $B(x, y)$, которые нужно вывести на ближний и тыльный экраны, требуется решить систему интегральных уравнений (3), (6) с ограничениями

$$0 \leq F(x, y) \leq 255; \quad 0 \leq B(x, y) \leq 255, \quad (7)$$

которые связаны с техническим диапазоном интенсивности изображений на ЖК экранах. Эти уравнения являются уравнениями Фредгольма первого рода с ядром типа свертки. Если изображения являются цветными, то систему уравнений (3), (6) с ограничениями (7) следует решить для каждого из цветов.

Метод решения системы уравнений, основанный на использовании нейросети. Учитывая, что изображение на ЖК панелях состоит из отдельных пикселей, для решения системы уравнений удобно применить искусственную нейронную сеть. Каждый *нейрон* моделирует свой пиксель, поэтому нейросеть будет состоять из двух слоев, соответствующих передней F и задней B панелям. Нейроны каждого слоя расположены в узлах двумерной равномерной сетки и в пределах каждого слоя не связаны между собой (рис. 5).

Между каждым нейроном слоя F существуют постоянные связи только с теми нейронами слоя B , которые участвуют в формировании изображений для левого и правого глаза. Эти нейроны соответствуют пикселям, лежащим на прямых (2), (5) и попадающим в зону влияния ядра (4), как это показано на рис. 5. При этом связи определяются ядром:

$$W_{kl} = K(x - x_k, y - y_l),$$

где x_k, y_l — координаты пикселя слоя B , через который проходят прямые (2), (5); x, y — координаты остальных пикселей слоя B .

Настраиваемыми величинами являются *уровни возбуждения нейронов*. На вход каждого нейрона тыльного слоя подается постоянный возбуждающий сигнал, соответствующий свету заднего осветителя S , который после последовательного прохождения тыльного и переднего слоев формирует на выходе нейросети изображения N_L и N_R для левого и правого глаза.

Формирование изображений нейросетью описывается следующими соотношениями:

$$N_L(i, j) = F(i, j) \times \sum_k \sum_l W_{kl} B(i + \Delta_L(i) - k, j + \Delta_L(j) - l);$$

$$N_R(i, j) = F(i, j) \times \sum_k \sum_l W_{kl} B(i + \Delta_R(i) - k, j + \Delta_R(j) - l),$$

где Δ_L, Δ_R — смещения индексов для левого и правого изображений, которые определяются геометрическими параметрами системы: расстоянием между глазами наблюдателя $x_L + x_R$, расстоянием от наблюдателя до дисплея D и расстоянием между панелями Δ .

Для того чтобы избавиться от ограничений (7), введем переменные f и b такие, что

$$F(i, j) = \frac{255}{1 + \exp(-f(i, j))};$$

$$B(i, j) = \frac{255}{1 + \exp(-b(i, j))}.$$

В результате выходные сигналы нейронов будут находиться в диапазоне $[0, 255]$, а переменные f и b могут принимать любые значения на числовой оси.

Для настройки нейросети будем минимизировать следующий функционал:

$$\Phi(f, b) = \sum_i \sum_j \left(L(i, j) - N_L(i, j) \right)^2 + \sum_i \sum_j \left(R(i, j) - N_R(i, j) \right)^2 + \alpha \sum_i \sum_j \left(f(i, j)^2 + b(i, j)^2 \right). \quad (8)$$

Первые два слагаемых этого функционала описывают разницу между выходом нейросети и изображениями, которые должны попасть соответственно в левый и правый глаз. Последнее слагаемое является регуляризирующей добавкой; оно позволяет выделить из множества

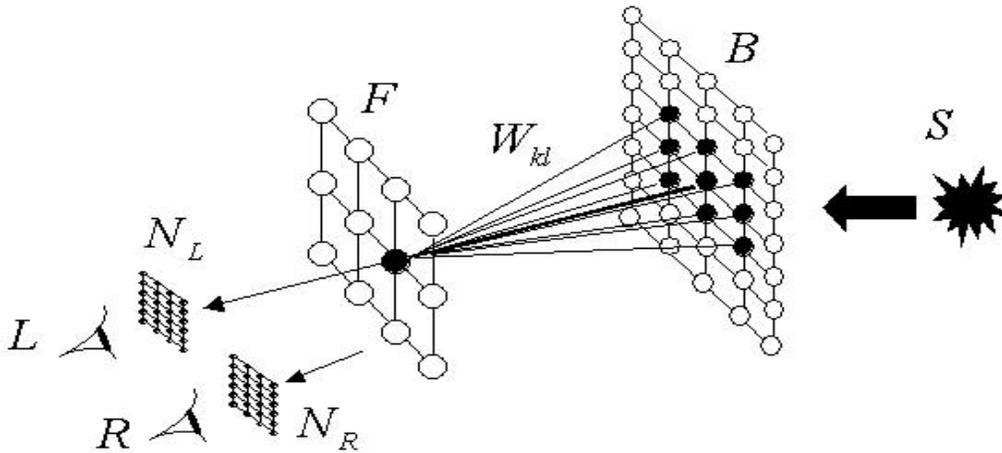


Рис. 5. Архитектура нейросети для нахождения видеоизображения на стереодисплее

решений единственное, имеющее наименьшую норму. Здесь α — параметр регуляризации, численные эксперименты показали, что для него оптимальным является значение из интервала $10^{-3} \div 10^{-2}$.

Для минимизации функционала (8) наиболее эффективным оказался метод градиентного спуска с инерционным членом:

$$f(t+1) = f(t) - \eta \frac{\partial \Phi(t)}{\partial f} - \mu \frac{\partial \Phi(t-1)}{\partial f};$$

$$b(t+1) = b(t) - \eta \frac{\partial \Phi(t)}{\partial b} - \mu \frac{\partial \Phi(t-1)}{\partial b}.$$

Здесь η, μ — параметры, характеризующие скорость настройки; t — фиктивный параметр, означающий номер итерации.

Итерационный процесс уверенно сходится при любом начальном распределении f и b , однако скорость сходимости существенно повышается, если в качестве начального приближения для изображений на экранах выбирается, например, квадратный корень из интенсивности изображений стереопары: $F(i, j) = \sqrt{L(i, j)}$; $B(i, j) = \sqrt{R(i, j)}$. Это соответствует следующему начальному распределению для f и b :

$$f(i, j) = -\ln \left(\frac{255}{\sqrt{L(i, j)}} - 1 \right);$$

$$b(i, j) = -\ln \left(\frac{255}{\sqrt{R(i, j)}} - 1 \right).$$

Метод решения, основанный на преобразовании Фурье. В приближении отсутствия диффузора можно предложить эффектив-

ную процедуру вычисления искомых изображений для переднего и тыльного экранов, основанную на быстром преобразовании Фурье. Это приближение имеет практический смысл, если черная маска, отделяющая одну ячейку от другой, занимает существенно меньшую площадь, чем ЖК ячейки.

При отсутствии диффузора система уравнений (3), (6) примет вид

$$L(x, y) = F(x, y) B(ax - \Delta x_L, ay);$$

$$R(x, y) = F(x, y) B(ax - \Delta x_R, ay). \quad (9)$$

Эта система двумерных уравнений распадается на набор одномерных систем для каждой строки изображений при фиксированных значениях y , поэтому в данном разделе будем опускать второй аргумент.

Прологарифмируем уравнения (9), в результате получим следующую систему:

$$l(x) = f(x) + b(ax - \Delta x_L);$$

$$r(x) = f(x) + b(ax - \Delta x_R),$$

где $l(x) = \ln L(x)$; $r(x) = \ln R(x)$; $f(x) = \ln F(x)$; $b(x) = \ln B(x)$. Применим к этой системе преобразование Фурье:

$$F_\omega(l(x)) = F_\omega(f(x)) + \frac{1}{a} e^{i\omega \Delta x_L/a} F_\omega \left(\frac{b(x)}{a} \right);$$

$$F_\omega(r(x)) = F_\omega(f(x)) + \frac{1}{a} e^{i\omega \Delta x_R/a} F_\omega \left(\frac{b(x)}{a} \right).$$

Обозначив $V = \left(F_\omega(f(x)), F_\omega \left(\frac{b(x)}{a} \right) \right)^T$, $U = \left(F_\omega(l(x)), F_\omega(r(x)) \right)^T$, запишем систему в следующем виде:

$$A(\omega)V = U,$$

где A — квадратная матрица размером 2×2 :

$$A = \begin{pmatrix} 1 & \frac{1}{a}e^{i\omega\Delta x_L/a} \\ 1 & \frac{1}{a}e^{i\omega\Delta x_R/a} \end{pmatrix}.$$

Определитель этой матрицы равен $\frac{2i}{a} \times \sin\left(\omega\frac{\Delta x_R}{a}\right)$, он обращается в нуль по крайней мере в точке $\omega = 0$. Для получения устойчивого решения можно воспользоваться схемой регуляризации Тихонова [6]:

$$U_\alpha = F^{-1}(A^*A + \alpha E)^{-1}V,$$

где A^* — эрмитово-сопряженная с A матрица; α — параметр регуляризации.

Результаты численного моделирования. Примеры изображений, которые нужно вывести на ближний и тыльный экраны, приведены на рис. 6. В верхнем ряду (см. рис. 6,а,б) находится исходная стереопара, т. е. изображения, которые должны видеть соответственно левый и правый глаз. В среднем ряду (см. рис. 6,в,г) приведены изображения, полученные с помощью метода, основанного на искусственной нейросети. И наконец, в нижнем ряду (см. рис. 6,д,е) приведены изображения, рассчитанные с помощью метода регуляризации Тихонова. Поскольку в последнем случае предполагается отсутствие диффузора, изображения получились более темными по сравнению с предыдущим случаем.

Математическая модель проекционной системы

Выберем систему координат с началом в центре экрана, причем горизонтальная ось Ox и вертикальная ось Oy лежат в плоскости экрана, а ось Oz направлена в сторону размещения проекторов. Разместим проекторы на одном уровне по высоте. В этом случае можно зафиксировать значение y и решать задачу для каждой строки изображения.

Рассмотрим произвольную точку экрана x_C , которую освещают два проектора P_L и P_R , как это показано на рис. 2. Пусть луч левого проектора P_L , попадающий в эту точку, имеет интенсивность I_L^P , а луч правого проектора P_R , попа-

дающий в эту точку, — интенсивность I_R^P . Тогда для левого глаза наблюдателя, находящегося в точке L , отраженное от этой точки изображение будет иметь интенсивность

$$I_L(x) = I_L^P K(\alpha_1) + I_R^P K(\beta_1),$$

где α_1 — угол между лучами $P_L x_C$ и $x_C L$; β_1 — угол между лучами $P_R x_C$ и $x_C R$. Аналогичным образом записывается интенсивность изображения, отраженного от точки x_C и попадающего в правый глаз:

$$I_R(x) = I_L^P K(\alpha_2) + I_R^P K(\beta_2).$$

Задача состоит в нахождении таких изображений I_L^P и I_R^P , выводимых на проекторы, чтобы левый и правый глаз видели заданные изображения D_L и D_R с учетом ограничений

$$0 \leq I_L^P \leq 255; \quad 0 \leq I_R^P \leq 255. \quad (10)$$

Ограничения обусловлены тем, что интенсивность пикселей проекторов должна находиться в диапазоне $[0, 255]$.

Метод нахождения изображений, основанный на использовании нейросети. С учетом того, что как изображения, выводимые на проекторы, так и изображения стереопары состоят из отдельных пикселей, т. е. имеют дискретную структуру, для построения искомого изображения был разработан метод, основанный на использовании нейросети специальной архитектуры. Предлагаемая нейросеть содержит три слоя нейронов, первый из которых моделирует возбуждающие сигналы, а остальные моделируют пиксели изображения. Архитектура нейросети изображена на рис. 7.

Нейроны первого слоя, обозначенного буквой x , моделируют возбуждающие сигналы. Они соединены постоянными связями со значениями, равными единице, с соответствующими нейронами второго слоя. Нейроны второго слоя y соответствуют одной строке пикселей левого и правого проекторов. Нейроны третьего слоя z соответствуют пикселям изображения, попадающего в левый и правый глаз наблюдателя. При этом для определенности левая половина нейронов второго слоя соответствует левому проектору, правая половина — правому проектору. Аналогично, левая половина нейронов третьего слоя соответствует изображению для левого глаза, правая — для правого глаза, поэтому число



Рис. 6. Результаты численного эксперимента для стереодисплея: *a, б* — исходная стереопара; *в, г* — изображения, выводимые на ближний и тыльный экраны, рассчитанные с использованием нейросети; *д, е* — изображения для ближнего и тыльного экранов, полученные с помощью метода регуляризации Тихонова

нейронов в каждом слое равно $2n$, где n — число пикселей в одной строке изображения.

Связи между нейронами второго и третьего слоев W_{ij} являются постоянными, они моделируют геометрические свойства оптической схемы. То есть между нейроном y_i и z_j существует связь, если в j -й пиксель изображения, видимого

наблюдателем, попадает изображение, прошедшее через i -й пиксель проектора. Значение каждой связи определяется углом между источником света и наблюдателем, т. е. функцией рассеяния (1). С выхода нейронов третьего слоя снимаются изображения I_L и I_R соответственно для левого и правого глаза.

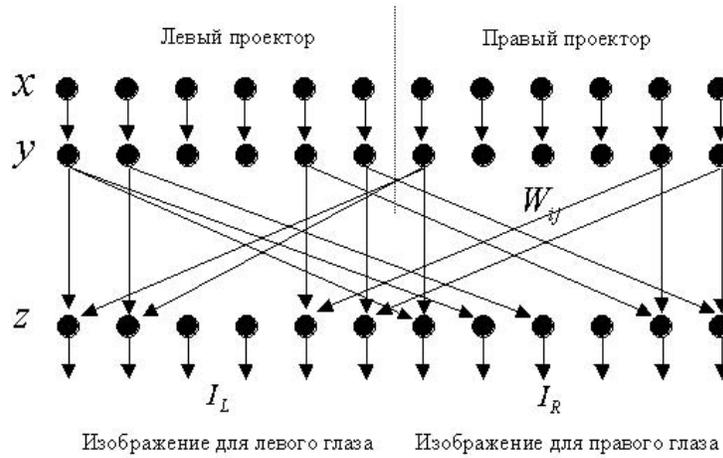


Рис. 7. Архитектура нейросети, используемой для нахождения видеоизображений на проекторах

Настраиваемыми величинами являются уровни возбуждения нейронов первого слоя x_i . В процессе настройки минимизируется отклонение выходных изображений I_L и I_R от изображений на заданной стереопаре D_L и D_R . При этом на выходе нейронов второго слоя формируются искомые изображения I_L^P и I_R^P , выводимые на проекторы.

Прямое распространение сигнала по нейросети определяется следующими формулами:

$$y_i = \frac{255}{1 + e^{-x_i}}, \quad i = 1, \dots, 2n; \quad (11)$$

$$z_i = \sum_{j=1}^{2n} W_{ij} y_j, \quad i = 1, \dots, 2n.$$

В качестве ошибки нейросети предлагается следующая величина:

$$E = \frac{1}{2} \sum_{i=1}^n (D_{L,i} - I_{L,i})^2 + \frac{1}{2} \sum_{i=1}^n (D_{R,i} - I_{R,i})^2 + \alpha \sum_{i=1}^n (I_{L,i}^2 + I_{R,i}^2), \quad (12)$$

где $I_{L,i} = z_i$, $i = 1, \dots, n$; $I_{R,i} = z_{n+i}$, $i = 1, \dots, n$; α — параметр регуляризации.

Первые два слагаемых функционала (12) описывают разницу между выходом нейросети и изображениями, которые должны попасть соответственно в левый и правый глаз. Последнее слагаемое является регуляризирующей добавкой; оно позволяет выделить из множества решений единственное, имеющее наименьшую норму [6]. Неоднозначность в решении появляется потому, что из-за геометрических свойств

оптической схемы часть нейронов второго слоя связана более чем с двумя нейронами третьего слоя, а это означает, что один и тот же пиксель проектора должен участвовать в формировании двух различных пикселей изображения.

Использование сигмоидной функции в формуле (11) позволяет избавиться от ограничений (10), т. е. дает возможность проводить минимизацию ошибки нейросети на всей числовой оси.

Для настройки нейросети наиболее эффективным оказался метод градиентного спуска с инерционным членом:

$$x_i(t+1) = x_i(t) - \eta \frac{\partial E(t)}{\partial x_i} - \mu \frac{\partial E(t-1)}{\partial x_i},$$

$$i = 1, \dots, 2n.$$

Здесь η , μ — параметры, характеризующие скорость настройки; t — фиктивный параметр, означающий номер итерации. Итерационный процесс уверенно сходится при любом начальном распределении, однако скорость сходимости существенно повышается, если в качестве начального приближения для x_i выбираются значения, при которых выходы нейронов второго слоя соответствуют исходной стереопаре:

$$x_i = -\ln \left(\frac{255}{D_L} - 1 \right), \quad i = 1, \dots, n;$$

$$x_i = -\ln \left(\frac{255}{D_R} - 1 \right), \quad i = n+1, \dots, 2n.$$

Результаты численного эксперимента. На рис. 8 представлены результаты численных расчетов изображений, выводимых на два про-

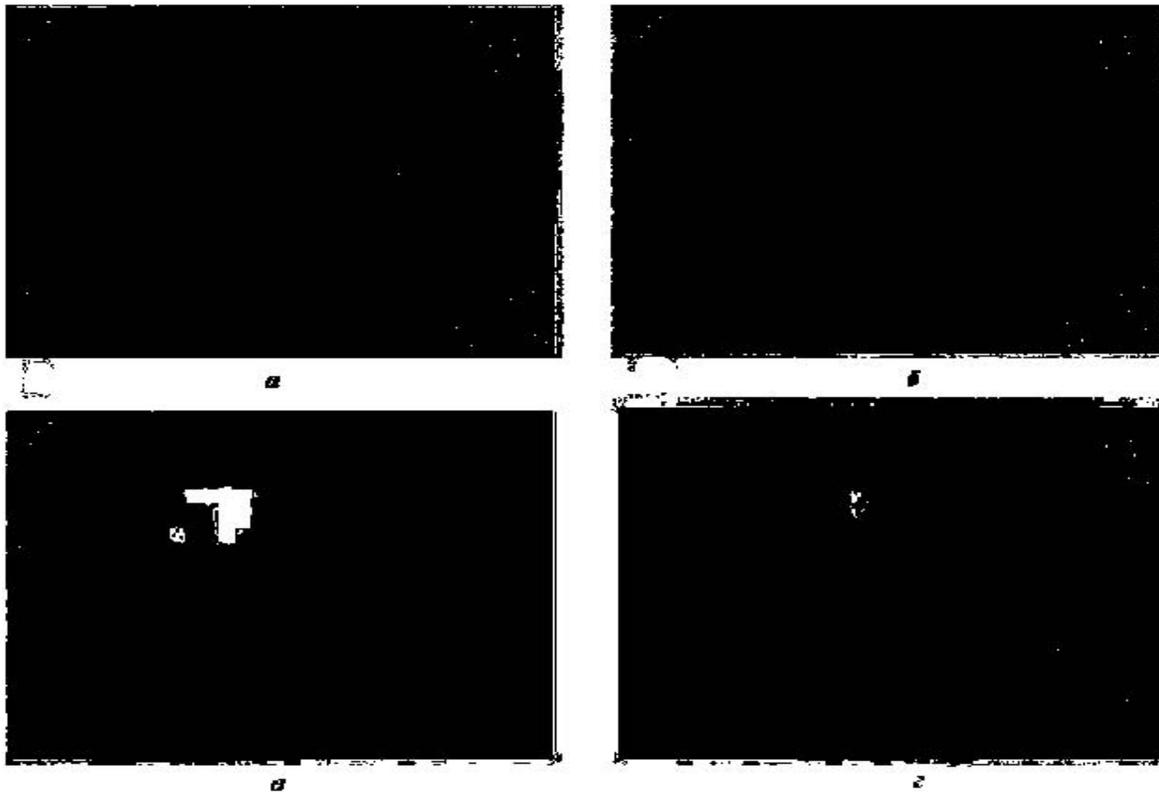


Рис. 8. Результаты численных экспериментов: *a, б* — исходная стереопара; *в, г* — изображения, которые нужно вывести соответственно на левый и правый проекторы

ектора. На рис. 8,*a,б* представлена исходная стереопара, т. е. изображения, которые должны видеть левый и правый глаз. На рис. 8,*в,г* — изображения, которые получились в результате расчетов: изображение слева следует вывести на левый проектор, изображение справа — на правый. При расчетах были выбраны следующие параметры: $\eta = 0,05$; $\mu = 0,1$; $\alpha = 0,1$. Процесс настройки сошелся за 45 итераций, при этом максимальная ошибка нейросети, отнесенная к одному пикселю, не превышала 8 градаций, что составило 3 %.

Заключение

Предложенные автостереоскопические устройства были реализованы "в железе": компания NeurOK Optics LLC выпустила два типа дисплеев — Smartron с двумя ЖК панелями и Tritron с тремя ЖК панелями. Как эти дисплеи, так и проекционная система, содержащая четыре проектора, неоднократно демонстрировались на международных выставках, таких как CeBit, Comdex и IfoComm.

Список литературы

1. Лукьяница А., Путлин А. Способ воспроизведения изображения объекта // Российское агентство по патентам и товарным знакам, 1999. RU 2158949 C1.
2. Lukyanitsa A. A., Putilin A. N. Visualization of three-dimensional images and multiaspect imaging. United States Patent 6985290, 2004.
3. Lukyanitsa A. A., Putilin A. N. Stereodisplay with neural network image processing // Proc. SPIE. 2002. Vol. 4660. P. 207—211.
4. Lukyanitsa A. A., Putilin A. N. Three-dimensional image projection employing retro-reflective screens. United States Patent 6843564, 2005.
5. Калинин А., Воронин П. Перегнать, не догоняя // М.: Компьютерра. 2006. №13. С. 31—33.
6. Тихонов А. Н., Арсенин В. Я. Методы решения некорректных задач. М.: Наука, 1979.