

Roman STAROSOLSKI

OBRAZY STEREOSKOPOWE

Streszczenie. Artykuł przedstawia aspekty postrzegania głębi przez człowieka oraz charakteryzuje komputerowe metody generowania obrazów stereoskopowych. Zostały także przedstawione doświadczenia z implementacją wybranych technik, propozycje zwiększenia realizmu i ułatwienia obserwacji obrazu oraz prosty algorytm analizy autostereogramów.

STEREOSCOPIIC IMAGES

Summary. In this paper human depth perception is described. Computer methods of generating stereoscopic images are reviewed and implementation of selected ones is presented. Improvements enhancing image realism and making viewing easier are proposed. A simple algorithm of analyzing autostereograms is also presented.

STEREOSKOPISCHE BILDER

Zusammenfassung. In diesem Artikel ist menschliche Tiefperzeption beschrieben und computerische Generierungsmethoden der stereoskopischen Bilder charakterisiert. Da sind auch Versuche mit der Implementation den einigen Methoden, Vorschläge von der Rälismuserhöhung und Beobachtungserleichterung des Bildes, sowie einfacher Analysealgorithmus der Autostereogrammen vorgestellt.

1. Postrzeganie głębi przez człowieka

Do rozpoznania kształtu i wzajemnego położenia obiektów mózg wykorzystuje zarówno różnice w obrazach docierających z obu oczu, jak i szczególne cechy charakterystyczne pojedynczego obrazu. Czynniki postrzegania głębi można podzielić na psychologiczne i fizjologiczne. Wśród psychologicznych wymienia się:

- wielkości znanych obiektów — jeżeli drzewo i człowiek na obrazie mają taką samą wysokość, to człowiek znajduje się bliżej, a drzewo dalej,
- perspektywa liniowa — objekty o stałych rozmiarach w miarę oddalania się mają coraz mniejsze rozmiary kątowe, np. tory kolejowe wydają się zbiegać w jeden punkt,
- perspektywa atmosferyczna — powietrze nie jest w stu procentach bezbarwne i przezroczyste, np. w górach odległe szczyty wydają się bardziej niebieskie,
- przesłanianie — jeżeli pierwszy obiekt częściowo przesłania drugi, to musi znajdować się przed nim,
- rozkład światła i cieni — przedmioty rzucają cienie, same się częściowo w tych cieniach znajdują, dlatego od strony padania promieni świetlnych przedmioty wypukłe są jaśniejsze,
- paralaksa ruchu — podczas jazdy pociągiem słupy trakcji poruszają się szybciej, daleko mijane budynki miasta wolniej, a słońce stoi w miejscu,
- zmiany wielkości obiektów — jeżeli obraz samochodu się zwiększa to znaczy, że samochód zbliża się do obserwatora, ale jeżeli bańka mydlana na końcu słomki się zwiększa, to być może ktoś ją nadmuchuje.

Czynniki fizjologiczne to:

- akomodacja — zmiana ogniskowych soczewek tak, by ostro dojrzeć obserwowany obiekt, objekty widziane nieostro muszą się znajdować w innej odległości od głowy niż te widziane ostro,
- zbieżność — takie wzajemne nachylenie oczu, by obraz przedmiotu znalazł się na plamce żółtej obu oczu,
- nierównoodległość siatkówkowa — różnica między obrazami tego samego obiektu na siatkówkach wynikająca z różnicy położenia oczu i ich kierunków.

Pomimo faktu, że tylko dwa ostatnie czynniki mają charakter obuoczny, to właśnie w tej grupie znajduje się czynnik uznawany [7] za decydujący o percepcji głębi — nierównoodległość siatkówkowa, czyli różne przesunięcia poziome występujące między obiektami na obu obrazach. Tradycyjnie odnosi się ją do dobrze zdefiniowanych krawędzi i brzegów obiektów. Należy jednak zauważyć [2], że cienie, które dają wyraźne zmiany jasności powierzchni, jakby dodatkowe krawędzie, zdecydowanie ułatwiają rozpoznawanie kształtów, a zwłaszcza krzywizn powierzchni.

Według Hodgesa i McWorthera [4], gdzie porównywano wrażenia z obserwacji mono- i stereoskopowych obrazów typu: cieniowane wypełnione powierzchnie, klatka drucziana (czyli obraz składający się z samych krawędzi obiektów) i klatka drucziana z usuniętymi niewidocznymi krawędziami, oczywiście najbardziej realistyczne obrazy to stereoskopowe z cieniowaniem. Po nich następują monoskopowe z cieniowaniem, stereoskopowe klatki drucziane z usuniętymi liniami niewidocznymi i stereoskopowe klatki drucziane. Natomiast kształt i położenie obiektów najłatwiej jest odtworzyć z obrazów stereoskopowych typu klatka drucziana bez usuwania niewidocznych linii.

Generowane komputerowo prymitywne obrazy stereoskopowe charakteryzują się nieskończoną głębią ostrości. Takie obrazy mogą być odbierane jako ponadnaturalnie ostre i niezwykle wyraziste. Dzieje się tak, ponieważ mózg ludzki, przyzwyczajony do koncentrowania się na ostro widzianym fragmencie obrazu, nie jest w stanie zignorować, nie występujących w obrazie, fragmentów nieostrych i musi się dopiero nauczyć wybierania punktu skupienia uwagi spośród wielokrotnie większej niż w rzeczywistości przestrzeni.

Generując obraz stereoskopowy można się postarać, by nie przysparzać dodatkowej pracy dla mózgu i zasymulować głębi ostrości. Do realizacji tego celu potrzebna jest, niestety nie zawsze dostępna, informacja — co jest aktualnie najważniejszym obiektem na obrazie. Inny problem, dużą złożoność czasową generacji nieostrego obrazu, można ominąć stosując prostsze od zmniejszania ostrości techniki różnicowania obiektu zainteresowania od reszty obrazu (np. obiekt w centrum uwagi rysować jaśniejszymi liniami).

Znając punkt skupienia uwagi można spróbować rozwiązać pojawiający się problem, gdy na ekranie widnieje obraz obiektu rzeczywiście znajdującego się za lub przed nim: zbieżność oczu ustawiona jest na jedną odległość, akomodacja na inną. Mózg wykorzystuje nierównoodległość siatkówkową do określenia odległości względnych, a rozmiary znanych obiektów do oszacowania odległości bezwzględnych. Jeżeliby więc przesunąć i przeskalować scenę tak, aby punkt skupienia uwagi znalazł się na płaszczyźnie ekranu i obiekty na ekranie nie zmieniły rozmiarów, to zbieżność będzie zgodna z akomodacją, a odległości względne i bezwzględne będą mogły być nadal poprawnie wydobywane z obrazu.

Precyzja rozpoznawania głębi położenia obiektów w punkcie przecięcia się osi optycznych oczu jest nadspodziewanie duża — według [3] wynosi kilka sekund kątowych, a więc dziesięć razy więcej niż, szacowana na jedną minutę kątową, rozdzielczość pojedynczego oka. Odpowiada to rozpoznaniu z dystansu pół metra przesunięcia w głąb o jedną piątą milimetra. Takim wymaganiom nie może sprostać typowy wyświetlacz komputerowy, gdyż przesunięcie w przeciwnie strony, o pojedynczy piksel rastra (0.25mm), obrazów dla obu oczu na ekranie oglądanym z odległości pół metra da wrażenie zmiany położenia o prawie cztery milimetry. Rozdzielczość ta maleje jednak wykładniczo w miarę oddalania się od tego punktu (w odległości piętnastu minut kątowych od punktu skupienia uwagi wynosi już tylko 2-3 minuty kątowe), a powyżej odległości 20 minut kątowych jest stała — jest to argument przemawia-

jący za dokładnym przedstawieniem obiektu w centrum zainteresowania obserwatora. Mniej istotne fragmenty obrazu mogą być niezbyt wiernie przedstawiane, ponieważ szacuje się [1], że z około miliarda bitów docierających w ciągu sekundy do oczu do świadomości przedostaje się zaledwie sto.

2. Metody otrzymywania stereogramów

2.1. Hełm stereoskopowy

Hełm stereoskopowy jest zespołem układów optycznych umieszczanych na głowie obserwatora i przekazujących bezpośrednio do jego oczu obraz z dwóch miniaturowych wyświetlaczy. Hełm jest połączony z komputerem za pomocą kabla transmitującego obrazy do wyświetlaczy. Głównym obszarem stosowania hełmów stereoskopowych są rozbudowane systemy rzeczywistości wirtualnej. Wtedy obserwator ma do dyspozycji także specjalny manipulator pozwalający na nawigowanie w trójwymiarowej przestrzeni. Niektóre tego typu systemy umożliwiają dodatkowo śledzenie pozycji głowy przez komputer generujący obrazy — dzięki temu wrażenie głębi nie ulega zakłóceniom mimo poruszania głową.

Hełm ma wiele zalet nieosiągalnych za pomocą innych metod. Najważniejszymi są: wysoka jakość kolorowych obrazów docierających niezależnie do każdego oka oraz poprawna praca mimo przechylenia i poruszania głowy. Pozytywną cechą jest także ograniczenie pola widzenia obserwatora do generowanego obrazu — hełm eliminuje wszelkie zewnętrzne bodźce wzrokowe (a zazwyczaj także i słuchowe) dając obserwatorowi wrażenie pełnego zanurzenia się w symulowanej przestrzeni.

Tego typu systemy wymagają nie tylko specjalizowanych wyświetlaczy, ale czasami również specjalizowanych stacji graficznych generujących obraz. Co za tym idzie, ich podstawową wadą jest bardzo wysoka cena. Z oczywistych przyczyn liczba obserwatorów jest ograniczona do jednego.

2.2. Metody z podziałem czasu

Ideą metod z tej grupy jest przekazywanie człowiekowi jednocześnie tylko jednego obrazu na przemian przez lewe i prawe oko. Obraz jest wyświetlany na monitorze i obserwator musi mieć wzrok skierowany na monitor — nie ma więc takiej swobody jak w przypadku hełmu. Dzięki bezwładności oka, tak jak w przypadku zwyczajnych monitorów, użytkownik nie zauważy migotania, jeżeli tylko obrazy będą przekazywane z wystarczającą częstotliwością. Częstotliwość ta musi być wyższa niż w przypadku zwykłych monitorów, ponieważ obraz docierający do pojedynczego oka ma ponad dwa razy niższy współczynnik wypełnienia

(stosunek długości czasu wyświetlania obrazu do odstępu między kolejnymi obrazami). Praktycznie do wyświetlania obrazów metodą z podziałem czasu stosuje się zwyczajne monitory ekranowe pracujące w trybach z przeplotem.

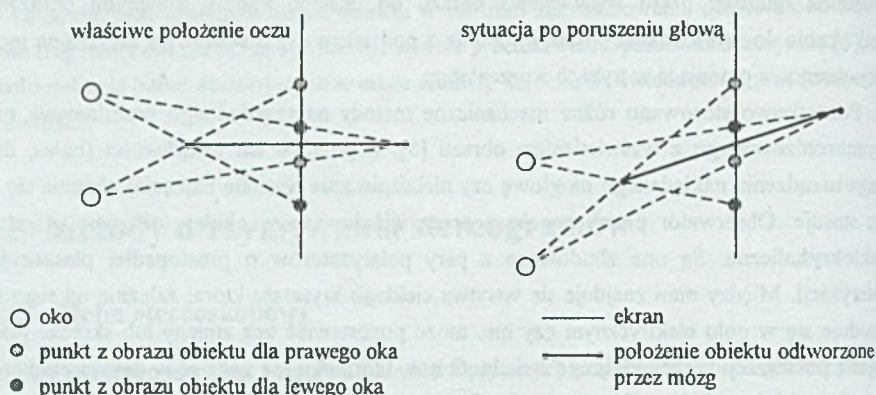
Początkowo stosowano różne mechaniczne metody naprzemiennego przesłania oczu zsynchronizowanego z wyświetlaniem obrazu [5]. Z powodu ich uciążliwości (hałas, duża waga urządzenia nakładanego na głowę czy niebezpiecznie wysokie napięcia) obecnie się ich nie stosuje. Obserwator patrzy na ekran przez wbudowane w okulary aktywne przesłony ciekłokrystaliczne. Są one zbudowane z pary polaryzatorów o prostopadłej płaszczyźnie polaryzacji. Między nimi znajduje się warstwa ciekłego kryształu która, zależnie od tego czy znajduje się w polu elektrycznym czy nie, może pozostawiać bez zmiany lub skręcać płaszczyznę polaryzacji przechodzącego światła. O tym, które oko ma widzieć w danym momencie ekran, decyduje układ monitorujący sygnał video docierający do wyświetlacza. Na podstawie tego sygnału układ rozpoznaje, który półobraz będzie wyświetlany i informację tę przekazuje do okularów drogą połączenia elektrycznego lub za pomocą światła podczerwonego.

Obraz generowany z podziałem czasu ma, z punktu widzenia pojedynczego oka, dwukrotnie mniejszą częstotliwość i dwukrotnie mniejszą rozdzielczość pionową. Jednak, jak zauważono w [4], subiektywna jakość obrazu zależy silniej od częstotliwości, ponieważ użytkownicy wolą obraz 120Hz z przeplotem (dla pojedynczego oka 60Hz i połowa rozdzielczości pionowej) od obrazu 60Hz (odpowiednio 30Hz i pełna rozdzielczość).

Jeżeli obraz ma być oglądany przez dużą grupę osób, to można zastosować pewną modyfikację wyżej opisanej metody. Obserwatorów wyposaża się w pasywne okulary, zamiast całych przesłon są w nie wbudowane tylko pojedyncze tylne polaryzatory o prostopadłych płaszczyznach polaryzacji. Warstwa ciekłego kryształu i przedni polaryzator znajdują się na ekranie — razem z okularami tworzą układ dopuszczający obraz jednocześnie tylko do jednego oka. Rozwiązanie z pasywnymi okularami jest dla jednego obserwatora droższe, ale dodawanie kolejnych obserwatorów jest już dużo tańsze, co więcej — pojedyncze pasywne polaryzatory są wygodniejsze i lżejsze od wymagających zasilania i sterowania przesłon.

Metody prezentowania stereogramów na zwyczajnym monitorze techniką z podziałem czasu, jak również komputerowe anaglify, zazwyczaj nie umożliwiają śledzenia położenia oczu. Najlepiej nadają się więc do prezentacji obrazów stereoskopowych wtedy, gdy można przewidzieć położenie głowy obserwatora i założyć, że nie będzie on nią poruszał. Skutki poruszania głową, gdy jej pozycja nie jest śledzona przez komputer, obrazuje rys. 1.

Taka sytuacja może mieć miejsce w sali projekcyjnej, w przypadku gdy stereogram jest prezentowany z dużej odległości większej liczbie osób siedzących nieruchomo w fotelach (niewielkie zmiany położenia głowy są nieistotne w porównaniu do odległości między głową a ekranem). Obserwator może także powstrzymać się od poruszania głową świadomie i znając ograniczenia takich metod prezentacji trójwymiarowej sceny nauczyć się, że jeżeli chce obejrzeć scenę z lewej strony, to zamiast przesuwania głowy w lewo powinien wydać



Rys. 1. Zmiana kształtu obiektu po zmianie położenia oczu
 Fig. 1. Change of shape after eye-movement

odpowiednie polecenie programowi generującemu obraz. Pozornie wydaje się, że jest to stawianie zbyt dużych wymagań obserwatorowi, lecz tak właśnie postępują osoby obserwujące zwykłe obrazy na ekranie komputera i można sądzić, że po pewnym czasie przyzwyczajania się będą także tak postępowały z obrazami stereoskopowymi.

Zaletą rozwiązań z podziałem czasu jest wielokrotnie niższa od helmu cena, jak również możliwość realizacji na odpowiednio rozbudowanym standardowym sprzęcie. Obrazy generowane tą metodą może jednocześnie oglądać spore grono osób.

Wadą technik z podziałem czasu są, poza obniżoną częstotliwością przekazywania obrazów, przede wszystkim niedoskonałości techniczne samej metody. Występują w niej bowiem przesłuchy między obrazami. Każde oko widzi oprócz swojego obrazu także słaby drugi obraz. Przyczyn tego stanu rzeczy jest wiele:

- dynamiczna niedoskonałość przesłon — ciekły kryształ nie reaguje natychmiastowo,
- statyczna niedoskonałość przesłon — przesłona otwarta przepuszcza tylko 33% światła, zamknięta blokuje jedynie 98%,
- poświata fosforu ekranu — poprzedni obraz nie zdąży całkowicie zgasnąć przed wyświetleniem następnego.

Mimo że wady metody są nie do wyeliminowania na obecnym etapie rozwoju techniki, systemy z podziałem czasu z powodzeniem stosuje się w wysokiej klasy, czyli drogich, systemach CAD.

2.3. Komputerowe anaglify

Anaglif to dwubarwny rysunek wykonany zazwyczaj za pomocą kolorów czerwonego i niebieskiego. Powstaje poprzez nadrukowanie na siebie dwu rzutów płaskich figury: rzutów w kierunkach odpowiadających położeniu oczu u przeciętnego człowieka. Do oglądania anaglify potrzebne są kolorowe okulary. Każde oko ogląda rysunek przez filtr przepuszczający tylko obraz przeznaczony dla niego i w ten sposób do mózgu docierają dwa różne obrazy. Po pewnym czasie, na skutek adaptacji chromatycznej wzroku (jest to efekt częściowej utraty wrażliwości oka na dominujący kolor — dzięki temu identyfikujemy barwy przedmiotów oświetlonych w różny sposób, np. światłem słonecznym i sztucznym) człowiek przestaje zauważać barwy filtrów i odbiera obraz jako bezbarwny w stopniach jasności. Anaglify to najpowszechniej znane obrazy stereoskopowe. Obecnie wykorzystuje się je na przykład w podręcznikach geometrii wykreślnej, gdzie bez nich odtworzenie przestrzennego kształtu ze skomplikowanego płaskiego rzutu byłoby trudne, a z ich pomocą odbywa się podświadomie.

Uwagę zwraca fakt, że zdecydowanie najniższą czułość wykazują pręciki wrażliwe na kolor niebieski [1]. Być może więc da się polepszyć jakość odbioru anaglifów poprzez zastosowanie zamiast barw czerwonej i niebieskiej barw czerwonej i zielonej oraz odpowiednich filtrów na oczy.

Komputerowa wersja tradycyjnych anaglifów ma w stosunku do anaglifów drukowanych w książkach wiele zalet:

- mogą być generowane dynamicznie przez komputer,
- można je dostosować do indywidualnego odbiorcy — na przykład o innym niż przeciętny rozstawie oczu.

Komputerowe anaglify można, w przeciwieństwie do opisywanych poprzednio metod, generować w oparciu o standardowy i powszechnie dostępny sprzęt. Do ich realizacji oprócz kolorowego monitora konieczne są jedynie okulary z kolorowymi filtrami.

Barwne filtry na ogół nie są najwyższej jakości — nie blokują w 100% innych barw, często nie mają dokładnie takiej samej barwy jak luminofor ekranu, a i oczy człowieka nie są idealnie takie same. Za pomocą operacji na palecie barw wyświetlacza będzie można, przynajmniej częściowo, skompensować ich niedoskonałości kosztem co najwyżej niewielkiego pogorszenia kontrastu obrazu.

Podstawową zaletą anaglifów jest dostępność sprzętu koniecznego do ich generacji, a także koszty zwiększenia liczby obserwatorów, które są minimalne.

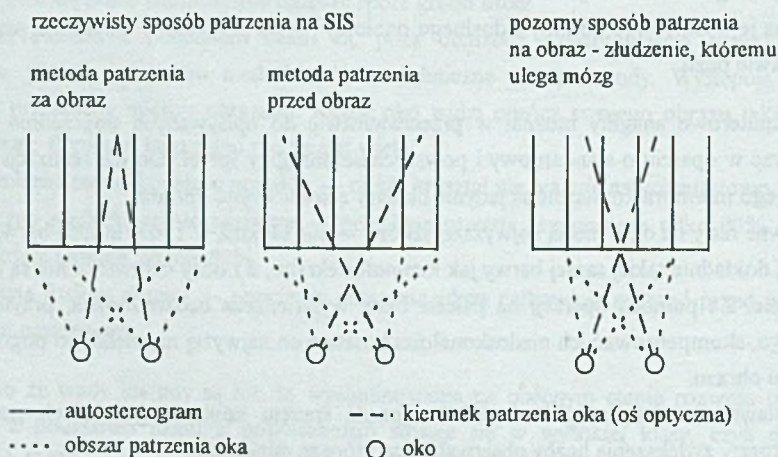
Największą wadą jest niemożność uzyskania koloru tą techniką. Za każdym razem trzeba też przyzwyczajać wzrok do ignorowania kolorów filtrów, gdyż tuż po założeniu okularów do obserwatora docierają obie rywalizujące ze sobą barwy.

2.4. Autostereogramy

Określane także jako SIS (ang. Single Image Stereogram) obrazy przeznaczone są do obserwowania bez żadnego dodatkowego oprzyrządowania. Generowane za pomocą komputerów zazwyczaj przedstawiają statyczne sceny i dlatego równie dobrze jak na ekranie prezentują się wydrukowane na kartce papieru.

Autostereogramy zdecydowanie różnią się od obrazów uzyskiwanych za pomocą pozostałych metod. Stymulują tylko jeden z czynników percepcji głębi — mechanizm nierównoodległości siatkówkowej. W każdej innej metodzie tworzy się dwa oddzielne rzuty sceny w kierunku oczu. Tutaj obraz powstaje nie z dowolnej sceny, ale z tzw. topograficznej mapy wysokości — płaskiego obrazu (rys. 4a), w którym jasność poszczególnych punktów określa względną wysokość powierzchni sceny. W konsekwencji nie da się przedstawić przesłaniania obiektów, można jedynie uzyskać wrażenie różnej względnej wysokości fragmentów obrazu. Jeżeli jednak zmiana poziomu wysokości jest gwałtowna, to mózg może zinterpretować taki „oderwany” kawałek powierzchni jako osobny obiekt.

Do zrozumienia metody oglądania autostereogramów przydatna jest znajomość ich struktury. Składają się one z pionowych pasów zawierających albo pozornie przypadkowo porzucane punkty przypominające biały szum, albo jakiś powtarzający się wzorec. Autostereogram można interpretować jako nałożone na siebie dwa obrazy (przesunięte o szerokość pasa i wygenerowane w ten sposób, że nakładające się fragmenty są identyczne).



Rys. 2. Metody i mechanizm oglądania autostereogramów
Fig. 2. Methods and mechanism of watching autostereograms

Jeżeli obserwator nie będzie normalnie obserwował kartki, a spojrzy na nią tak, by jednym okiem widzieć pasy od pierwszego do przedostatniego, a drugim od drugiego do ostatniego, to do mózgu dotrą dwa różne obrazy — spełniony zostanie więc warunek niezbędny do uzyskania wrażenia głębi.

Można zatem oglądać autostereogramy na dwa sposoby (rys. 2):

1. Patrząc daleko za kartkę (tzw. metoda patrzenia za obraz) — wtedy lewe oko będzie widziało pas numer 1, a w obrazie prawego oka będzie mu odpowiadał pas numer 2. Większość autostereogramów jest przeznaczona do oglądania w ten właśnie sposób.
2. Patrząc na wymaginowany punkt znajdujący się przed kartką (tzw. metoda patrzenia przed obraz) — oko lewe patrzy na pasy od 2 do n, a prawe na pasy od 1 do n-1.

Jeżeli do autostereogramu wygenerowanego dla jednej z metod spróbuje się zastosować drugą, to uzyska się obraz odwrócony, to co było przed, znajdzie się za.

Dla ułatwienia na niektórych SIS (rys. 3) narysowane są dwa, odległe od siebie o szerokość pasa, znaczniki. Na obraz należy patrzeć tak, by widzieć trzy znaczniki (nie dwa ani nie cztery). Sam obraz jest stworzony w ten sposób, że punkty w obrazie z jednego oka pokrywają się z punktami dla drugiego jeśli obszar, na którym są położone, ma znajdować się na poziomie kartki. Jeżeli ma być wyżej, to punkty w obrazie dla prawego oka są przesunięte na lewo w stosunku do obrazu dla lewego oka. Im płaszczyzna sceny położona jest dalej od kartki, tym przesunięcie względne jest większe. Fragmenty znajdujące się za kartką przesuwane są w przeciwną stronę.

Do mózgu docierają dwa różne obrazy, są one traktowane jako dwa obrazy tego samego obiektu. Ponieważ obrazy te różnią się między sobą, to na zasadzie mechanizmu nierównoodległości siatkówkowej (wyłącznie na zasadzie tego mechanizmu) mózg interpretuje przesunięcia fragmentów jednego obrazu względem odpowiednich obszarów drugiego (obrazu) jako różne wysokości względne powierzchni, której dwa obrazy do niego dotarły.

Oglądanie autostereogramów nie zawsze jest sprawą prostą. Zazwyczaj, by się tego nauczyć, potrzeba mniej niż godzinę. Część osób nie jest w stanie patrzeć ze zbieżnością oczu ustawioną na jedną odległość, a akomodacją na inną i widzi obraz nieostry, inni w ogóle nie potrafią świadomie sterować zbieżnością oczu. Według dokumentacji do jednego z shareware'owych programów do przygotowywania autostereogramów wynika, że 10%, a z moich spostrzeżeń, że aż 30%, ludzi nie potrafi tego dokonać.

Zastosowanie autostereogramów jest ograniczone, były głównie wykorzystywane do badań nad ludzkim zmysłem wzroku. Obecnie używa się ich w reklamie, mogą być także traktowane jako specyficzna forma sztuki.

3. Implementacja wybranych technik

3.1. Komputerowe anaglify

Generowanie anaglifów na ekranie komputera daje dobre rezultaty w porównaniu do anaglifów drukowanych. Najważniejszą różnicą jest dynamika obrazu komputerowego — łatwiejsze do oglądania są sceny ruchome, przy czym duże znaczenie ma płynność animacji, a co za tym idzie — i szybkość komputera generującego ją. Implementacja metody komputerowego anaglifu jest kompromisem pomiędzy szybkością generowania kolejnych scen a ich realizmem. W celu otrzymania realistycznych anaglifów, pozwalających na skuteczne odtworzenie geometrii oryginalnej sceny, program generuje dynamicznie obraz typu klatka drucziana bez usuniętych niewidocznych linii i zwiększa jego realizm poprzez symulację głębi ostrości i płynną animację.

Jako komputerowa wersja drukowanych, np. w książkach do geometrii wykreślnej, anaglifów spisuje się bardzo dobrze — dzięki modyfikacjom zwiększającym realizm prezentowanej sceny i dzięki możliwości dostosowania animacji do pozycji oraz charakterystyki okularów indywidualnego użytkownika.

Podstawowym wymogiem programu było działanie w czasie rzeczywistym oraz założenie, że obraz dla obu oczu musi pojawiać się na ekranie jednocześnie, jak również sprawdzenie wcześniej opisanych koncepcji:

- dostosowanie obrazu do konkretnego obserwatora i jego okularów, czyli także kompensowanie niedoskonałości okularów i wzroku samego człowieka za pomocą modyfikacji barw obrazu prezentowanego na ekranie,
- możliwie niedroga, pod względem złożoności czasowej, symulacja głębi ostrości dla zwiększenia realizmu anaglifu oraz wyróżnienia istotnego fragmentu sceny, jeżeli można przewidzieć, gdzie spogląda obserwator,
- takie przetransformowanie sceny, by zbieżność i akomodacja oczu ustawione były na tę samą odległość (na odległość ekranu) — tak jak i w przypadku symulacji głębi ostrości wymagana jest znajomość oczekiwanego punktu skupienia uwagi obserwatora.

Anaglify generowane są w pamięci i kopiowane na ekran. Rysowanie kolejnych linii odbywa się za pomocą operacji sumy logicznej (OR) z tłem, dzięki czemu jeżeli linia niebieska przetnie czerwoną, to w punkcie przecięcia kolor piksela będzie sumą kolorów linii, czyli będzie fioletowy. W trybach z paletą numer barwy piksela jest indeksem do palety aktualnie dostępnych barw. Tak więc aby logiczne sumowanie barw działało, należy odpowiednio wypełnić paletę.

Rysowanie operacją OR powoduje, że jeżeli przetną się dwie linie tej samej barwy, ale o różnej jasności, to wypadkowa barwa w miejscu przecięcia będzie nie ciemniejsza od

jaśniejszej z barw i nie jaśniejsza od sumy jasności obu barw. Ten efekt nie jest dużym problemem w typowych obrazach typu klatka druciana. Może jednak się nim stać w bardzo złożonych scenach, albo w przypadku rysowania z wypełnianiem powierzchni — nowa barwa nie przykrywa starej (kolejność rysowania jest nieistotna, ponieważ operacja OR jest przemienne).

Kompensowanie niedoskonałości okularów odbywa się przez indywidualną regulację jasności i kontrastu barw niebieskiej i czerwonej. To rozwiązanie jest dobre, gdy wadą filtrów w okularach jest różna ich jasność. Jeżeli jeden z filtrów częściowo przepuszcza światło barwy drugiego, to regulacja jasności i kontrastu nie rozwiąże tego problemu.

Symulacja głębi ostrości odbywa się poprzez przedstawianie nieostrych fragmentów sceny za pomocą grubszych i ciemniejszych kresek. Mimo wielkiej prostoty zastosowanej metody jest ona skuteczna, zwiększa bowiem realizm obrazu i kieruje uwagę obserwatora na jaśniejszy i wyraźniejszy fragment sceny.

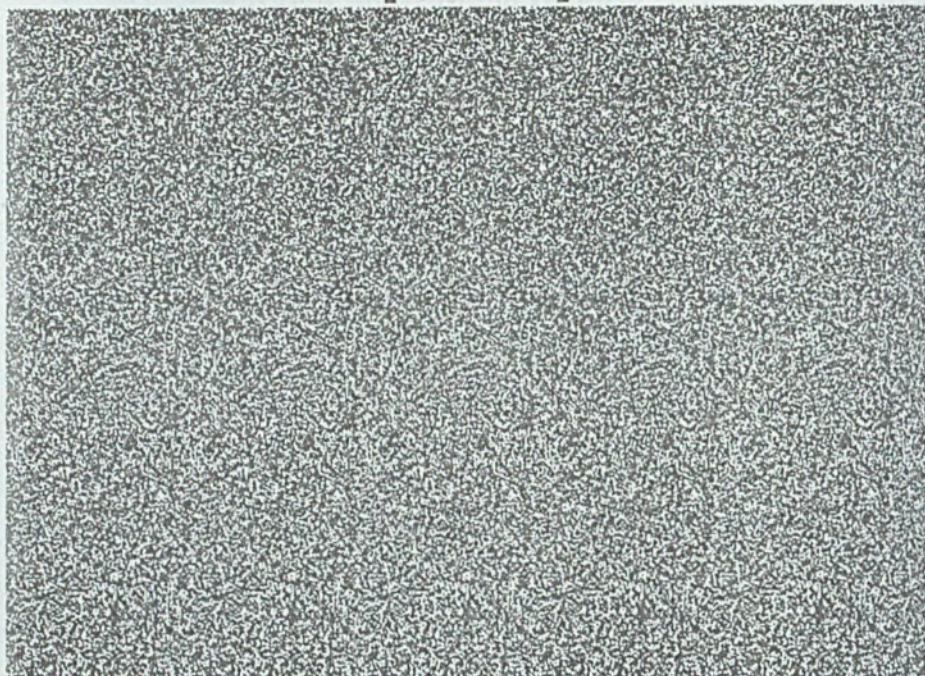
Generowanie obrazu, dla którego zbieżność oczu jest równa ich akomodacji (czyli punkt skupienia uwagi obserwatora znajduje się na płaszczyźnie ekranu), ułatwia obserwowanie anaglifów. Pomimo że odległości bezwzględne są teraz oceniane bez udziału zbieżności oczu, w animowanych scenach obserwator nadal poprawnie je szacuje.

Jeżeli założymy, że obserwator spogląda w dany punkt przestrzeni i dla tego punktu będziemy symulować głębię ostrości obrazu oraz transformować scenę uzyskując zgodność zbieżności oczu z akomodacją, to skłonimy obserwatora do skupienia się na wybranym przez nas fragmencie sceny (jeżeli wcześniej jego uwaga była skierowana gdzie indziej).

Mankamentem zrealizowanej implementacji okazał się nie rozwiązany problem ustalania początkowych wartości położenia głowy obserwatora i rozstawu jego oczu. Przyjęta w programie metoda pozwala użytkownikowi na zmianę standardowych wartości proponowanych przez program. Początkowe wartości są ustalone dla przeciętnego przypadku i zawsze dają się tak zmienić, by polepszyć odbiór anaglifów. Dotyczy to zwłaszcza rozstawu oczu obserwatora. Zmniejszanie tej wartości oczywiście powoduje „spłaszczanie” sceny, a zwiększanie poprawia wrażenie głębi. Zbyt duża wartość powoduje, iż obserwator zaczyna widzieć odległe od płaszczyzny rzutowania fragmenty sceny podwójnie. Z czasem obserwator przyzwyczaja się i może akceptować większe wartości.

3.2. Autostereogramy

W przeciwieństwie do pozostałych metod uzyskiwania obrazów stereoskopowych autostereogramy generuje się na podstawie mapy wysokości. Jest to narysowany przez użytkownika obraz, w którym kolory punktów (zależnie od konkretnej implementacji liczy się numer koloru, jasność lub temperatura barwy) decydują o ich względnej wysokości.



Rys. 3. Autostereogram
Fig. 3. The autostereogram

Z gotowego autostereogramu mózg odtwarza poziomy wysokości poszczególnych elementów sceny tylko na podstawie nierównoodległości siatkówkowej punktów w obu obrazach. Krawędzie powstają na granicy obszarów o różnych wysokościach.

Cały autostereogram jest tworzony na podstawie wzorca, który w gotowym autostereogramie będzie stanowił pierwszy pas. Czasami można też spotkać autostereogramy, w których pierwotny wzorec jest umieszczony na środku. Wzorec zazwyczaj jest losowo wypełniony punktami czarno-białymi — autostereogram wygenerowany na podstawie takiego wzorca nazywany jest RDS (ang. Random Dot Stereogram). Spotyka się również autostereogramy z kolorowych punktów, jak również autostereogramy wygenerowane na podstawie wzorca zawierającego kolorowy, abstrakcyjny wzór.

Tutaj zostanie opisany najprostszy algorytm generacji autostereogramów, tj. algorytm ze spoglądaniem wstecz. Generuje on SIS wierszami, dzięki czemu ma minimalne wymagania pamięciowe (wystarczy jednocześnie przechowywać jeden wiersz wzorca i jeden mapy zamiast całego wzorca i całej mapy).

Znaczenie zmiennych i tablic występujących w algorytmie:

- wzorzec — tablica z elementami pojedynczego wiersza wzorca
- dł_wzorca — długość pojedynczego wiersza wzorca
- mapa — mapa wysokości pojedynczego wiersza sceny, jej elementy są liczbami całkowitymi ze znakiem. Zero oznacza poziom odniesienia, liczba dodatnia fragment sceny powyżej poziomu odniesienia, ujemna — poniżej
- dł_mapy — długość pojedynczego wiersza mapy wysokości
- aster — tablica, do której algorytm wygeneruje pojedynczy wiersz autostereogramu (o długości dł_mapy+dł_wzorca)
- il_wierszy — ilość wierszy w autostereogramie, mapie i wzorcu
- pm — aktualna pozycja w mapie wysokości
- pa — aktualnie generowana pozycja w autostereogramie
- p0 — zmienna pomocnicza, pozycja w już gotowym fragmencie autostereogramu, z której element należy skopiować do pozycji aktualnie generowanej
- w — zmienna pomocnicza, licznik wierszy

```
FOR w:=1 TO il_wierszy DO
  BEGIN
    wczytaj wiersz wzorca
    wczytaj wiersz sceny

    FOR pa:=1 TO dł_wzorca DO aster[pa]:=wzorzec[pa];

    FOR pm:=1 TO dł_mapy DO
      BEGIN
        pa:=pm+dł_wzorca;
        p0:=pm+mapa[pm];

        IF p0<1 OR p0>=pa
          THEN p0=dowolna wartość z przedziału <1, pa-1>

        aster[pa]:=aster[p0];
      END

    zapisz wiersz autostereogramu
  END
```

Na podstawie pojedynczych wierszy wzorca i mapy wysokości sceny generowany jest pojedynczy wiersz autostereogramu. Operacja ta powtarzana jest dla wszystkich wierszy.

Na początek wiersza autostereogramu kopiowany jest wzorzec (o szerokości s_z_{wz} punktów) — staje się on pierwszym pasem dla oka lewego. Teraz na podstawie kolejnych punktów p mapy wysokości generujemy punkty $p+s_z_{wz}$ wiersza autostereogramu.

Aby uzyskać wrażenie wysokości względnej w generowanym punktu, trzeba na pozycję p wiersza obrazu dla oka prawego (czyli punkt $p+s_z_wz$ wiersza autostereogramu) skopiować punkt z pozycji $p+w$ wiersza obrazu dla oka lewego (punkt $p+w$ wiersza autostereogramu).

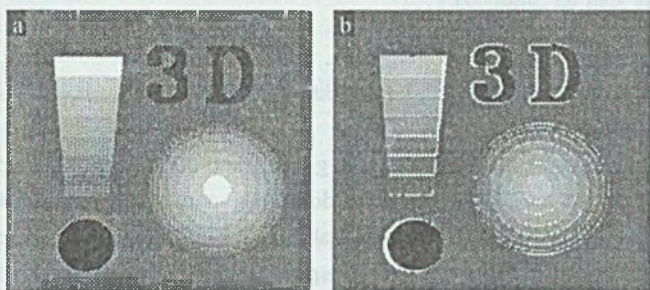
Jeżeli odległość punktu od poziomu odniesienia jest zbyt duża i trzeba by wziąć nie wygenerowany punkt z obrazu dla oka lewego, to można wybrać dowolny już wygenerowany punkt. Taką sytuację należy traktować jako błąd mapy wysokości.

Praktycznie najlepiej odbierane są autostereogramy, w których szerokość pasów mieści się w granicach od jednej do trzech czwartych odległości między oczami. Aby były czytelne, największa odległość punktów mapy wysokości od poziomu odniesienia nie powinna przekraczać $1/3$ szerokości wzorca, a poszczególne obiekty nie mogą mieć zbyt małych rozmiarów poziomych

3.3. Analizator autostereogramów

Odtworzenie sceny z gotowego autostereogramu można zrealizować naśladowując człowieka. Pierwszą czynnością jest ustalenie, przynajmniej w przybliżeniu, szerokości pasów, z których składa się autostereogram. W tym celu należy znaleźć w wierszu autostereogramu fragment powtarzający się i zapamiętać, jako offset bazowy, odległość między początkami powtórzeń. Zaczynając od minimalnej wartości offsetu bazowego, należy szukać identycznych fragmentów przesuniętych o jeden piksel. W ten sposób znajdzie się szerokość pojedynczego pasa, a nie jego wielokrotność. Aby uniknąć przypadkowej zgodności pikseli, szukany fragment powinien być jak największy, np. jedna czwarta długości wiersza. Jeżeli nie uda się znaleźć dużego fragmentu, to można całą operację powtórzyć po zmniejszeniu go o połowę. Jeżeli jednak po wielokrotnym zmniejszeniu długość szukanego fragmentu stanie się mniejsza od zadanego minimum, to poszukiwania należy przerwać, bo natrafi się na przypadkowe podobieństwa.

Następnie można przystąpić do dekodowania kolejnych wierszy. W każdym wierszu trzeba znaleźć fragmenty powtarzające się, przesunięte o offset bazowy. Podobnie jak poprzednio, nie mogą być one mniejsze od zadanego minimum. Znalezione podobieństwa to fragmenty sceny znajdujące się na poziomie odniesienia. Teraz można szukać fragmentów położonych poniżej i powyżej. Robi się to szukając powtarzających się fragmentów przesuniętych o coraz większą odległość od offsetu bazowego. Zakres tych poszukiwań w zrealizowanej implementacji jest ograniczony maksymalnie do 120 (co wynika z przyjętego formatu odtwarzanej sceny — 1 bajt na piksel), zazwyczaj jednak wystarcza mniejszy zakres — zadany zakres poszukiwań ma wpływ na prędkość działania programu. Poszukiwania są prowadzone jednocześnie poniżej i powyżej offsetu bazowego, ponieważ równie prawdopodobne jest znalezienie obiektu poniżej, jak i powyżej poziomu odniesienia. Jeżeli w znalezionym fragmencie część pikseli będzie już wcześniej zdekodowana, to bliższa offsetowi bazowemu odległość ma priorytet, dalszą należy uznać za przypadkową zbieżność.



Rys. 4. Oryginalna scena (a) i jej wygląd po odtworzeniu z autostereogramu (b)
 Fig. 4. Source scene (a) and scene as recovered from autostereogram (b)

Dla typowych autostereogramów powyższy algorytm daje obraz wyraźnie zaszumiony. Skuteczne okazało się wprowadzenie niewielkiej modyfikacji algorytmu: sprawdzanie jednocześnie, zamiast jednej, trzech kolejnych linii. Ta modyfikacja powoduje, że krawędzie poziome w wynikowym obrazie oznaczone są tak jak fragmenty nie rozpoznane i mają grubość dwu linii. W typowym RDS składającym się z dużych płaskich powierzchni udział krawędzi jest niewielki, a ich dodatkowe wyróżnienie nawet pomaga w znajdowaniu poszczególnych płaszczyzn. Nagłemu wzrostowi wysokości powierzchni sceny odpowiada grubsza (szerokość w poziomie) krawędź — rysunek jest bardziej plastyczny.

Minimalna długość powtarzających się fragmentów wiersza ma wpływ na zaszumienie obrazu i na minimalny poziomy rozmiar elementu sceny, który da się rozpoznać. Zwiększenie jej wartości usuwa z obrazu zakłócenia, ale powoduje także nierozpoznawanie przez program drobnych (pod względem szerokości) elementów sceny. Jej zmniejszanie oczywiście umożliwia rozpoznawanie drobniejszych detali ukrytych w autostereogramie, ale zbyt mała wartość powoduje wykrywanie w autostereogramach (zwłaszcza w dwubarwnych RDS) przypadkowych zbieżności utrudniając rozpoznanie tych właściwych.

Dla dekodowanego autostereogramu z rys. 3 udało się znaleźć optymalne wartości parametrów. Odtworzona scena (rys. 4b) jest kontrastowa i oprócz krawędzi płaszczyzn prawie nie ma w niej fragmentów nierozpoznanych, nie ma też zakłóceń.

Prędkość działania programu jest akceptowalna. Na komputerze 486 100 MHz analiza obrazu o wymiarach 640 na 480 trwa poniżej 3 sekund.

4. Uwagi

Spśród przedstawionych metod tylko anaglify i autostereogramy dają się zaimplementować na standardowym sprzęcie.

Za pomocą metody komputerowego anagliflu generowano obrazy typu klatka drucziana bez usuwania niewidocznych linii, pozwalające najdokładniej odtwarzać geometrię oryginalnej sceny, ale nie dające wystarczającego wrażenia realizmu. Przedstawione modyfikacje (po założeniu punktu skupienia uwagi obserwatora można generować obraz o zgodnej zbieżności oczu z ich akomodacją i symulować głębię ostrości) zwiększające realizm wrażeń i ułatwiające obserwację obrazów stereoskopowych mogą być równie dobrze stosowane w przypadku helmu stereoskopowego i metod z podziałem czasu.

Autostereogramy są zbyt niedoskonałe, nakładają bowiem duże ograniczenia na scenę i nie każdy potrafi je oglądać, by znalazły szersze praktyczne zastosowanie. Dzięki opieraniu się tylko na mechanizmie nierównoodległości siatkówkowej mogą być łatwo analizowane przez komputer. Implementacja algorytmu analizatora okazała się całkiem skuteczna. Algorytm analizy może być udoskonalony, tzn. obecna jego postać może wymagać ręcznego doboru parametrów analizy, nie rozpoznaje bardzo małych elementów, jak również może zostać wprowadzona w błąd przez obrazy zawierające w górnych wierszach coś innego niż sam autostereogram.

LITERATURA

- [1] Zabrodzki J.: Grafika komputerowa metody i narzędzia, WNT Warszawa 1994.
- [2] Puerta A. M.: The power of shadows: shadow stereopsis. Journal of the Optical Society of America [Optics and Image Science] Vol: 6, Iss: 2, Feb. 1989.
- [3] McKee, Levi, Browne: The imprecision of stereopsis. Vision Research Vol: 30, Iss: 11, 1990.
- [4] Hodges, McWhorter: Stereoscopic display for design visualization. Signal Processing: Image Communication Vol: 4, Iss: 1, Netherlands Nov. 1991.
- [5] Lipscomb J.S.: Experience with Stereoscopic Display Devices and Output Algorithms. Proceedings of the SPIE - The Internat. Society for Optical Engineering Vol: 1083, USA 1989.
- [6] Foley, Van Dam, Hughes: Computer Graphics Principles & Practice. Addison Wesley Reading Mass. 1992.
- [7] Southard D.A. Transformations for stereoscopic visual simulation. Computers & Graphics Vol: 16, Iss: 4, UK Winter 1992.

Wpłynęło do Redakcji 4 marca 1996 r.

Abstract

Human depth perception is a very complex problem. From quite a number of image depth cues (both monocular and binocular) couple major ones can be selected — i.e. retinal disparity is primary factor in relative distances judgement, absolute distances are estimated mainly through an angle extent of objects of known absolute size.

From reviewed methods of generation of stereoscopic images only anaglyphs and autostereograms do not require any additional equipment to be produced on a standard personal computer — experiments with implementation of those methods are described.

Anaglyphs (red-blue images watched with coloured glasses) gain much when generated online by a computer. Image can be adjusted to the actual observer (eye-distance, head-position, glasses characteristic). Some further improvements are examined, they require additional information — what is observer looking at (we can also make observer look at a selected focus point). Those improvements can be introduced to any but autostereogram method.

Following focus point (by moving objects to place focus point on a screen and scaling them to conserve object's angle extent) makes watching easier since accommodation is now consistent with convergence — relative and absolute distances can still be properly estimated through above-mentioned perception cues.

Wireframes without hidden lines removed are the best images to recover original scene geometry, but they lack on realism. Simulation of visual acuity depth, though very simple (out of focus objects are drawn with fatter and darker lines), improves image realism.

Autostereograms (fig. 3) are rather unusual pictures — inconvenient in viewing they do not require any equipment to be viewed and only shape is hidden in such stereogram.

A simple algorithm of generation is presented. Thanks to utilizing only a retinal disparity mechanism autostereograms (in opposition to other methods) are easily reconstructed into original scene. A pretty efficient (fig. 4) algorithm is also presented.