

О. Н. Романюк, О. В. Романюк, В. Г. Мальований

Вінницький національний технічний університет

Хмельницьке шосе, 95, 21021 Вінниця, Україна

Комбінування методів для прискорення зафарбовування

Розглянуто питання комбінованого використання в системах комп'ютерної графіки різних методів зафарбовування з метою підвищення продуктивності кінцевої візуалізації.

Ключові слова: зафарбовування, рендеринг, дифузна та спекулярна складові кольору.

Вступ

Формування просторових зображень є складним, багатоетапним обчислювальним процесом із використанням складних моделей і методів. При формуванні тривимірних зображень важливо не тільки достовірно відтворити форму об'єкта і його конструктивні особливості, але й правильно передати градації кольорів, що є визначальним при створенні ілюзії об'ємності тривимірного об'єкта на двовимірному екрані. Етап кінцевої візуалізації (рендеринг) [1] є найтрудомісткішим етапом графічного конвеєра, оскільки для кожної точки сцени залежно від розташування джерел світла та спостерігача, оптичних властивостей поверхонь, характеристик джерел світла та кривизни поверхні визначаються інтенсивності складових кольору точок зображення та їхні екранні координати. Процес візуалізації складних сцен з високою деталізацією поверхонь об'єктів може тривати десятки годин [2].

Оскільки етап кінцевої візуалізації є найтрудомісткішим етапом формування тривимірних зображень і становить 60–80 % від загального обсягу обчислень [3], доцільним є розробка методів і засобів, які б забезпечили суттєве спрощення процедур рендерингу як на програмному, так і апаратному рівнях, що дасть змогу досягти потрібної динаміки реалістичних зображень у реальному часі та значно скоротити тривалість візуалізації складних проектів.

Аналіз методів і постановка задачі

Найпростішим методом зафарбовування є однотонне тонування [1–4], згідно з яким для кожного плоского трикутника визначався вектор нормалі, а на його основі — колір. Складові трикутники об'єкта заповнюються одним кольором без його градації, що призводить до різкої зміни інтенсивності кольору на межах три-

кутників. Однотонне зафарбовування вимагає найменших обчислювальних затрат. Воно хоча й має низьку якість, але внесено до функцій Direct3D і підтримується тривимірними акселераторами.

До найпоширеніших методів зафарбовування відносять метод Гуро [1, 3]. Процес зафарбовування має такі стадії [3]: а) розраховують вектори нормалей до кожної грані; б) шляхом усереднення нормалей усіх граней, яким належить вершина, розраховують нормалі у вершинах трикутника (багатокутника); с) визначають інтенсивності кольору у вершинах багатокутника, використовуючи значення нормалей; д) зафарбовують ділянку, обмежену багатокутником, шляхом лінійної інтерполяції інтенсивностей кольору вздовж ребер, а потім і між ребрами вздовж кожного рядка растеризації. Останнім часом затінення за Гуро часто використовують як проміжну стадію [2] з інтерактивного формування 3D-зображення, покладаючи побудову повноцінної сцени на етап фінального рендерингу.

До недоліків методу Гуро можна віднести наступні: а) метод використовує для визначення інтенсивностей кольору лінійну інтерполяцію, в той час як дифузна та спекулярна складові кольору мають нелінійний характер зміни; б) не враховується локальна кривизна поверхні, оскільки вектори нормалей визначаються тільки для вершин трикутника; в) відблиски відтворюються тільки в разі, якщо вершини трикутників знаходяться в їхній зоні (відблиск, який не має спільних точок із вершинами трикутників, або розташований усередині трикутника, не буде сформовано); г) на межах двох трикутників проявляються смуги Маха [5], які пов'язані з літеральним гальмуванням на сітківці ока; д) має місце зміна інтенсивності кольору зображення від кадру до кадру, що виражається в миготінні, особливо відблисків, оскільки при формуванні динамічних зображень змінюється структура та положення вузлів триангуляційної мережі; е) наявність артефакту типу «зірка», який полягає у тому, що відблиск, який повинен мати форму еліпса, має форму зірки.

Хоча метод Гуро й продовжує широко використовуватися для формування тривимірних зображень у реальному часі, однак більш перспективним вважається метод Фонга [1–5], в якому замість значень інтенсивності кольору інтерполюються вектори нормалей, які потім використовуються у функції тонування для обчислення інтенсивності кольору кожного елемента зображення. Метод характеризується, порівняно з методом Гуро, значно більшими обчислювальними витратами, однак при цьому досягається краща локальна апроксимація кривизни поверхні і, як наслідок, отримують більш реалістичні зображення.

Згідно з методом Фонга [4, 6] для кожної точки поверхні визначають вектор нормалі \vec{N} до поверхні, вектори напрямку до джерела світла \vec{L} та спостерігача \vec{V} , за значенням яких розраховується інтенсивність кольору.

Дії, які виконуються для розрахунку вектора нормалі в довільній точці полігональної області, ідентичні до дій, які виконуються при зафарбовуванні за Гуро, з тією різницею, що:

1) замість інтенсивностей кольору у вершинах трикутника використовуються вектори нормалей, причому їхнє інтерполювання здійснюється вздовж трьох ортогональних напрямків;

2) після знаходження векторів нормалей виконують їхню нормалізацію;

3) за значеннями одиничних векторів нормалей в кожній точці полігональної області розраховують інтенсивність кольору. При цьому враховуються також оптичні властивості матеріалу поверхні і взаємне розташування джерела світла та спостерігача. Формули для розрахунку інтенсивності визначаються вибраною моделлю освітлення.

Аналіз [3] показав, що зафарбовування середньостатистичного трикутника методом Фонга в понад 50 раз триваліше порівняно з методом Гуро.

Оскільки метод Фонга, який забезпечує високу реалістичність формування графічних сцен, характеризується високою обчислювальною складністю і для багатьох застосувань не придатний для формування динамічних зображень, а значно простіший з обчислювальної точки зору метод Гуро не гарантує достатньої реалістичності, то виникає задача комбінованого використання цих методів для досягнення прийняттого компромісу між продуктивністю формування графічної сцени та якістю.

Розробка комбінованих методів зафарбовування

При зафарбовуванні тривимірних зображень важливо досягти високої продуктивності формування графічних сцен без суттєвої втрати якості. Методи зафарбовування, основані на інтерполюванні інтенсивностей кольору вздовж ребер і рядків растеризації трикутників, не забезпечують високої реалістичності відтворення об'єктів, оскільки не враховують кривизни поверхні та мають візуальні артефакти [4]. З іншої сторони, врахування кривизни поверхні за рахунок визначення векторів нормалей до її кожної точки з подальшим обчисленням інтенсивностей кольору є достатньо трудомісткою процедурою. У зв'язку з цим розробка підходів до комбінованого використання різних методів зафарбовування та моделей освітлення є актуальною, оскільки це забезпечить підвищення продуктивності.

Один із найбільш простих підходів до комбінованого використання різних методів зафарбовування полягає в тому, що з урахуванням кривизни поверхні, її матеріалу, позицій джерел світла та спостерігача, співвідношення складових кольору визначається його інтенсивність у виділених точках поверхні, а проміжні значення між ними розраховуються з використанням інтерполювання. При адитивному виборі розмірів цифрових сегментів забезпечується керований компроміс між швидкістю формування графічного зображення та його якістю.

Якщо розміри цифрових сегментів вибрати рівними степеню двійки, то обчислення значно спрощуються, оскільки ділення й множення на операнд, який дорівнює степені двійки, еквівалентні зсуву.

Розглянемо один із можливих підходів до комбінованого використання різних методів зафарбовування. Для цього розіб'ємо рядки растеризації трикутника на цифрові сегменти довжиною 2^k і знайдемо точні значення інтенсивностей кольору в його граничних точках. За отриманими значеннями шляхом лінійного інтерполювання обчислимо інтенсивності кольору у внутрішніх точках сегментів.

Якщо рядок растеризації має довжину $(x_B - x_A)$, то приріст вектора нормалей можна знайти за виразом $\Delta \vec{N}_{AB} = \frac{\vec{N}_B - \vec{N}_A}{x_B - x_A}$, де \vec{N}_A та \vec{N}_B — вектори нормалей від-

повідно в лівій і правій точках рядка растеризації. Для цифрових сегментів довжиною 2^k цей приріст дорівнює $\Delta\vec{N} = \Delta\vec{N}_{AB} \cdot 2^k = (\vec{N}_B - \vec{N}_A) \frac{2^k}{x_B - x_A}$. Шляхом кодової лінійної інтерполяції легко знайти вектори нормалей у граничних точках цифрових сегментів. Так, наприклад, у кінцевій правій точці цифрового i -го сегмента $\vec{N}_{p,i} = \vec{N}_{l,i} + \Delta\vec{N} \cdot 2^k$, де $\vec{N}_{l,i}$ — значення вектора нормалі в початковій точці сегмента. Очевидно, що множення в другому доданку можна замінити на нагромаджувальне додавання.

Розрахувавши вектор нормалі в точці i , можна знайти інтенсивність кольору в даній точці згідно з інтегральною формулою зафарбовування [1, 3, 6].

Якщо I_{p_i} та I_{l_i} — інтенсивності кольору в кінцевих точках i -го цифрового сегмента, розмір якого кратний степеню двійки, то в довільній точці v цього сегмента, використовуючи кодове лінійне інтерполювання, можна знайти

$$I_{v_i} = I_{l_i} + \frac{I_{p_i} - I_{l_i}}{2^k} \cdot v.$$

Останній цифровий сегмент у рядку растеризації трикутника в загальному випадку має розмір, який менший або дорівнює 2^k (рис. 1).

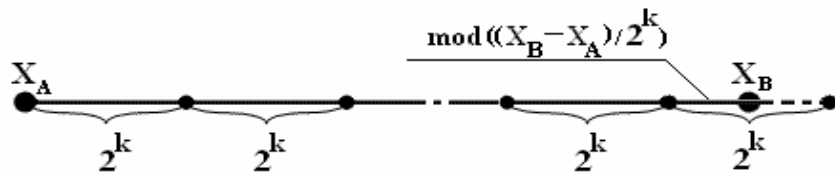


Рис. 1. Розбиття рядка растеризації на цифрові сегменти

У цьому випадку розмір останнього сегмента умовно вважають рівним степеню двійки, але розрахунки інтенсивностей кольору потокових точок усередині сегмента виконують тільки до кінцевої точки рядка растеризації. Це забезпечує однорідність обчислення для всього рядка растеризації.

Якщо $k = 1$, то цифровий сегмент містить 3 точки, причому значення інтенсивності кольору в середній точці дорівнює середньому значенню інтенсивностей кольору в граничних точках.

Оцінимо похибку, яка буде мати місце при розрахунку дифузної складової кольору. Інтенсивність дифузної складової кольору в крайніх лівій та правій точках сегмента знайдемо за формулами [1, 4]: $I_{d_{l_i}} = k_d \cdot I \cdot \vec{N}_{l_i} \cdot \vec{L}$; $I_{d_{p_i}} = k_d \cdot I \cdot \vec{N}_{p_i} \cdot \vec{L}$.

Найменший кут між вектором \vec{L} і одним із проміжних векторів буде мати місце у випадку, коли площина, в якій лежить вектор \vec{L} , буде перпендикулярною до площини, утвореною векторами \vec{N}_{l_i} і \vec{N}_{p_i} , а проекція вектора \vec{L} на зазначену площину збігається за напрямком з вектором $(\vec{N}_{l_i} + \vec{N}_{p_i})/2$ (максимально укороче-

ний порівняно з відповідним нормованим вектором).

Точне значення дифузної складової кольору в середній точці цифрового сегмента дорівнює [4]:

$$k_d \cdot I \cdot \frac{\vec{N}_{l_i} + \vec{N}_{p_i}}{\sqrt{2(1 + \cos \xi)}} \cdot \vec{L} = \frac{I_{d_{l_i}} + I_{d_{p_i}}}{\sqrt{2(1 + \cos \xi)}},$$

де ξ — кут між векторами \vec{N}_{l_i} , \vec{N}_{p_i} . При використанні кодового лінійного інтерполювання між значеннями $I_{d_{l_i}}$ і $I_{d_{p_i}}$ у середній точці цифрового сегмента отримаємо таке значення інтенсивності дифузної складової кольору — $(I_{d_{l_i}} + I_{d_{p_i}}) / 2 = Q_d / 2$.

Максимально можливе значення абсолютної похибки знаходимо за формулою:

$$\Delta_{\max} = \frac{I_{d_{l_i}} + I_{d_{p_i}}}{\sqrt{2(1 + \cos \xi)}} - \frac{I_{d_{l_i}} + I_{d_{p_i}}}{2} = Q_d \cdot \left(\frac{1}{\sqrt{2 \cdot (1 + \cos \xi)}} - \frac{1}{2} \right).$$

Графік зміни максимальної абсолютної похибки від кута ξ зображено на рис. 2.

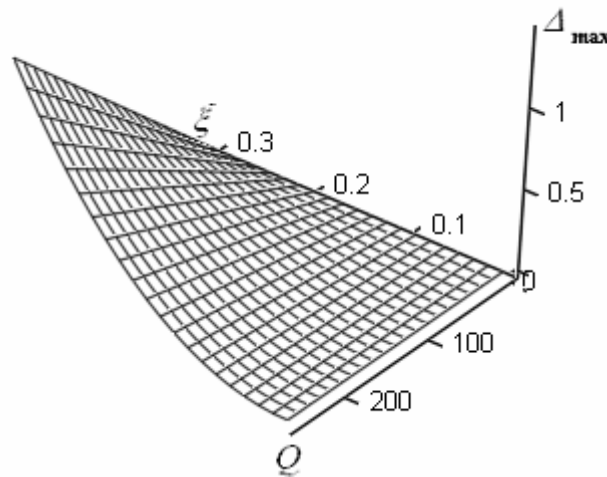


Рис. 2. Графік залежності Δ_{\max} від кута ξ

Знайдемо максимальне значення відносної похибки:

$$\delta_{\max} = \frac{(I_{d_{l_i}} + I_{d_{p_i}}) \cdot \left(\frac{1}{\sqrt{2 \cdot (1 + \cos \xi)}} - \frac{1}{2} \right)}{\frac{I_{d_{l_i}} + I_{d_{p_i}}}{\sqrt{2(1 + \cos \xi)}}} = \frac{2 - \sqrt{2 \cdot (1 + \cos \xi)}}{2} = 1 - \cos \frac{\xi}{2}.$$

На рис. 3 наведено графік зміни максимального значення відносної похибки δ від кута ξ .

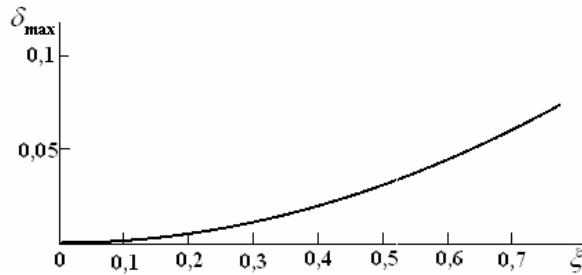


Рис. 3. Графік залежності δ_{\max} від кута ξ

При визначенні спекулярної складової кольору

$$I_{s l_i} = k_s \cdot I \cdot (\vec{N}_{l_i} \cdot \vec{H})^n, \quad I_{s p_i} = k_s \cdot I \cdot (\vec{N}_{p_i} \cdot \vec{H})^n,$$

де $\vec{H} = \frac{\vec{L} + \vec{H}}{|\vec{L} + \vec{H}|}$; n — коефіцієнт спекулярності поверхні.

Найменший кут між вектором \vec{H} і одним із проміжних векторів буде мати місце у випадку, коли площина, у якій лежить вектор \vec{H} , буде перпендикулярною до площини, утвореної векторами \vec{N}_{l_i} і \vec{N}_{p_i} , і його проекція збігається за напрямком із середнім вектором між зазначеними векторами, тому $\vec{N}_{l_i} \cdot \vec{H} = \vec{N}_{p_i} \cdot \vec{H} = Q$. При використанні кодової лінійної інтерполяції в середній точці цифрового сегмента отримуємо таке значення спекулярної складової кольору:

$$\frac{k_s \cdot I \cdot (\vec{N}_{l_i} \cdot \vec{H})^n + k_s \cdot I \cdot (\vec{N}_{p_i} \cdot \vec{H})^n}{2} = k_s \cdot I \cdot Q^n.$$

Точне значення спекулярної складової кольору в цій точці дорівнює:

$$k_s \cdot I \cdot \left(\frac{\vec{N}_{l_i} + \vec{N}_{p_i}}{\sqrt{2(1 + \cos \xi)}} \cdot \vec{H} \right)^n = k_s \cdot I \cdot \frac{(2 \cdot Q)^n}{(\sqrt{2(1 + \cos \xi)})^n}.$$

Максимальне можливе значення відносної похибки буде дорівнювати:

$$\delta_{\max} = \frac{k_s \cdot I \cdot \frac{(2 \cdot Q)^n}{(\sqrt{2(1 + \cos \xi)})^n} - k_s \cdot I \cdot Q^n}{k_s \cdot I \cdot \frac{(2 \cdot Q)^n}{(\sqrt{2(1 + \cos \xi)})^n}} = 1 - \frac{(\sqrt{2 \cdot (1 + \cos \xi)})^n}{2^n} = 1 - \cos^n \frac{\xi}{2}.$$

На рис. 4 наведено залежність δ_{\max} від n і ξ .

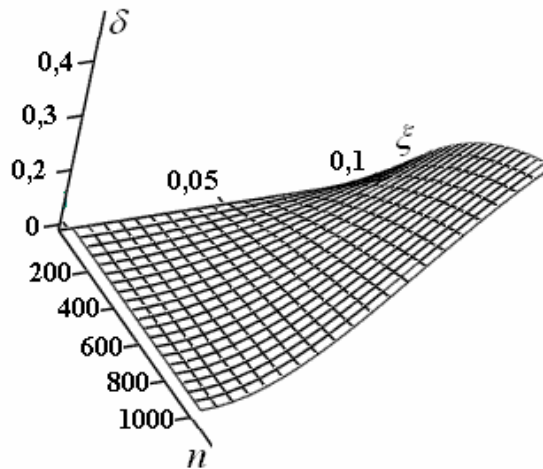


Рис. 4. Графік залежності δ_{\max} від кута ξ і коефіцієнта спекулярності n

Аналіз показав, що для найбільш поширених матеріалів ($n \leq 100$) $\delta_{\max} \leq 4\%$ при $\xi < 4^\circ$.

Якщо сегмент має розмір 2^k , то при його поділі навпіл отримуємо сегмент довжиною 2^{k-1} , причому середню точку сегмента буде розміщено у вузлі координатної решітки. Отримані сегменти знову мають розмір, кратний степеню двійки, тому при бінарному поділі всі точки отриманих сегментів буде розміщено у вузлах координатної решітки. Це дає можливість ефективно використати для зафарбовування принцип дихотомії [6]. При такому підході вектори нормалей у початковій і кінцевій точках сегмента знаходять за методом Фонга, а в усіх інших — із використанням принципу дихотомії, причому вектори нормалей визначаються точно.

Якщо відомо вектори нормалей \vec{N}_{l_i} і \vec{N}_{p_i} відповідно в лівій та правій граничних точках цифрового сегмента, то згідно з формулою [6]

$$\vec{N}_{(\frac{1}{2}^{i+1})} = \frac{\vec{N}_{l_i} + \vec{N}_{p_i}}{\sqrt{2 + z_{2^i}}}$$

У роботі [6] показано, що при використанні полінома першого степеня застосовується апроксимаційна формула

$$\frac{1}{\sqrt{2 + z_{2^i}}} \approx -0,07 \cdot z_{2^i} + 0,64$$

При цьому максимальна абсолютна похибка апроксимації не перевищує

0,0005, а відносна — 0,12 %. При використанні багаточлена другого степеня

$$\frac{1}{\sqrt{2+z_{2^i}}} \approx 0,014 \cdot z_{2^i}^2 - 0,119 \cdot z_{2^i} + 0,681.$$

Максимальна абсолютна похибка апроксимації не перевищує $2 \cdot 10^{-5}$, а відносна — 0,004 %.

Першу формулу доцільно використовувати для екранів із невеликою роздільною здатністю, для яких трикутники, які складають поверхню тривимірного об'єкта, мають незначні розміри

Дифузну складову кольору при послідовному поділі сегмента в сторону його лівої граничної точки можна знайти за формулою:

$$I_{d(\frac{1}{2^{j+1}})_i} = k_d \cdot I \cdot \vec{L} \frac{\vec{N}_i + \vec{N}_{(\frac{1}{2^j})_i}}{\sqrt{2+z_{2^j}}} = \frac{I_{d_i} + I_{d(\frac{1}{2^j})_i}}{\sqrt{2+z_{2^j}}}.$$

Вибір розміру цифрового сегмента визначається щільністю триангуляційної мережі, роздільною здатністю екранного простору, а також вимогами до реалістичності та продуктивності формування графічних сцен.

Аналіз показав, що при $k = 2$ досягається підвищення продуктивності при визначенні дифузної складової кольору в 1,3 рази, а при використанні апроксимації виразу $1/\sqrt{2+z_{2^n}}$ поліномом першої степені — в 1,8 рази. При $k = 3$ ці цифри складають відповідно 1,48 і 2,09.

При визначенні спекулярної складової кольору використовують формулу

$$I_{s(\frac{1}{2^{j+1}})_i} = k_s \cdot I \cdot (\vec{H} \frac{\vec{N}_i + \vec{N}_{(\frac{1}{2^j})_i}}{\sqrt{2+z_{2^j}}})^n.$$

Якщо у графічному процесорі реалізовано операцію піднесення до степеня, то можна застосовувати формулу, у якій використовуються скалярні операції

$$k_s \cdot I \left(\frac{I_{s_i} + I_{s(\frac{1}{2^j})_i}}{\sqrt{2+z_{2^j}}} \right)^n,$$

де $I_{s(\frac{1}{2^{j+1}})_i} = \vec{N}_i \cdot \vec{H}$, $I_{s(\frac{1}{2^j})_i} = \vec{N}_{(\frac{1}{2^j})_i} \cdot \vec{H}$.

Висновки

Розроблено метод зафарбовування, згідно з яким інтенсивності кольору всередині цифрових сегментів, на які розбито рядки растеризації, визначається із ви-

користанням кодового лінійного інтерполювання, а в кінцевих точках — із використанням векторів нормалі до джерела світла, поверхні та спостерігача, а також властивостей матеріалу. Вибір розмірів сегментів визначається адаптивно, залежно від необхідної точності відтворення кольору. Комбінування методів дозволяє підвищити продуктивність зафарбовування за рахунок використання спрощеної процедури зафарбовування для внутрішніх точок сегмента.

1. Херн Д. Компьютерная графика и стандарт OpenGL / Д. Херн, М. Бейкер. — М.: Издательский дом «Вильямс», 2005. — 1168 с.
2. Akenine-Moller T. Real-Time Rendering / T. Akenine-Moller, N. Hoffman., E. Haines. — Wellesley: A. Peters, 2007. — 1045 p.
3. Романюк О.Н. Високопродуктивні методи та засоби зафарбовування тривимірних графічних об'єктів. Монографія / О.Н. Романюк, А.В. Чорний. — Вінниця: УНІВЕСУМ-Вінниця, 2006. — 190 с.
4. Калютов А.В. Введение в фотореалистическую графику / А.В. Калютов. — СПб.: Политехника, 2006. — 118 с.
5. Роджерс Д. Математические основы машинной графики / Д. Роджерс, Дж. Адамс: Пер. с англ. — М.: Мир, 2001. — 604 с.
6. Романюк О.Н. Адаптивна нормалізація векторів нормалей при визначенні дифузної та спекулярної складових кольору / О.Н. Романюк, О.М. Мельников // Реєстрація, зберігання і оброб. даних. — 2006. — Т. 8, № 3. — С. 11–19.

Надійшла до редакції 30.11.2010