Slide 1:

Slide 2:

Nội dung trình bày hôm nay của nhóm em gồm có các phần như sau:

(đọc slide)

Việc phân chia trình bày chỉ mang tính tương đối, còn khối lượng công việc thực sự mà mỗi người đã thực hiện là như nhau trong quá trình thực hiện đề tài.

Slide 3:

Em xin được bắt đầu trình bày phần 1 – giới thiệu

Slide 4:

Đối với các nghiên cứu ở Trong lĩnh vực Computer Vision, mục tiêu là lấy thông tin từ các hình ảnh thu được để từ đó đáp ứng các nhu cầu thực tế.

Trong đó, độ sâu trong không gian của các điểm trên hình là 1 thông tin rất quan trọng. Có những vấn đề có thể sử dụng độ sâu để cho ra kết quả tốt hơn, đáng tin cậy hơn, như:

* Phát hiện đối tượng trong các bài toán hỗ trợ người tham gia giao thông, về việc nhận diện các mốc ranh giới trên đường, phát hiện chướng ngại vật, đặc biệt là người đi bộ dưới lòng đường
* Mô hình đối tượng, ứng dụng trong giải trí: các ứng dụng game điều khiển bằng cơ thể
* Một vấn đề khác rất cần thông tin độ sâu để giải quyết, đó là Free viewpoint, hệ thống sử dụng nhiều hình ảnh thu được từ camera thật để nội suy ra mọi góc nhìn của khung cảnh với góc nhìn của camera ảo.

Tuy nhiên, trong quá trình thu ảnh 2 chiều từ camera, thông tin về độ sâu bị mất. Cần phải tính toán để tìm lại thông tin này và ta có thể áp dụng nó vào 1 trong các bài toán đã đề cập.

Stereo Matching là bài toán tính toán thông tin độ sâu từ những ảnh 2D, xuất hiện tương đối lâu đời trong ngành Computer Vision và gần đây phát triển rất nhanh chóng, được rất nhiều sự quan tâm của các nhóm nghiên cứu và có nhiều công trình nổi bật.

Slide 5:

Do đó, nhóm cùng thầy GVHD đã lựa chọn tên đề tài là “Ứng dụng Stereo Matching trong phát hiện đối tượng” với mục tiêu:

* Khảo sát và đánh giá các công trình nghiên cứu về phương pháp Stereo Matching
* Từ khảo sát và đánh giá đó, lựa chọn nghiên cứu phù hợp để áp dụng thông tin độ sâu vào bài toán phát hiện đối tượng

Giới hạn của đề tài là những nghiên cứu mới và đạt kết quả tốt khi thực nghiệm và ứng dụng phát hiện đối tượng với các ảnh trong môi trường giao thông.

Slide 6:

Từ mục tiên và giới hạn của đề tài, nhóm đề ra nội dung cần phải thực hiện của đề tài, gồm:

* Tìm hiều những nghiên cứu trong giới hạn đề tài để thiết kế và xây dựng benchmark cho chúng
* Lựa chọn 1 nghiên cứu Stereo Matching phù hợp với bài toán phát hiện đối tượng trên đường, sau đó đề xuất và hiện thực phương pháp phát hiện đối tượng.
* Nội dung cuối cùng là thực nghiệm phương pháp đã đề xuất và đánh giá kết quả

Slide 7:

Tiếp theo, em xin trình bày các công trình liên quan.

Vì Stereo Matching là kiến thực nền tảng xuyên suốt nội dung đề tài nên em sẽ giới thiệu ngắn gọn khái niệm Stereo Matching trước khi đi vào những nghiên cứu liên quan trực tiếp.

Slide 8:

Đầu vào của 1 giải thuật Stereo Matching là 2 hay nhiều ảnh chụp cùng vị trí trong không gian với góc nhìn khác nhau.

Giải thuật Stereo Matching so sánh những điểm ảnh ở 2 ảnh chụp để tìm ra các cặp tương ứng và tính các giá trị disparity (độ lệch về vị trí) giữa các điểm tương ứng ở các bức ảnh. Từ những giá trị độ lệch đó, có thể tính toán được độ sâu của các điểm ảnh trong không gian.

Theo khảo sát và phân loại của D. Scharstein và R. Szeliski, một cách chung nhất, các giải thuật Stereo Matching được tiến hành theo quy trình gồm 4 bước:

1. Matching cost computation: tính chi phí so trùng

2. Cost (support) aggregation: tích hợp chi phí so trùng

3. Disparity computation / optimization: tính toán độ lệch

4. Disparity refinement: hiệu chỉnh độ lệch

Tuy nhiên, trong từng giải thuật cụ thể, các bước có thể được gộp (ví dụ bước 1 và bước 2) hoặc lược bớt (bước 4). Trong một số giải thuật, trình tự các bước có thể được thay đổi, VD: giá trị độ lệch được tính ở bước 3 có thể được sử dụng để hổ trợ bước 2, 1 phần hoặc toàn bộ 1 bước được thực hiện nhiều lần.

Tiếp sau đây, em sẽ trình bày sơ lượt ý tưởng thực hiện của mỗi bước, trước khi liệt kê nhanh 1 số cách tiếp cận cho từng bước.

Slide 9:

Như đã trình bày, đầu vào ở bước này là 2 ảnh Trái – Phải chụp cùng 1 khung cảnh trong không gian. Các ảnh đã được tiền xử lý nên chúng không lệch nhau theo phương đứng mà chỉ lệch theo phương ngang.

Gọi d là độ lệch tối đa thì tọa độ mỗi cặp điểm tương ứng sẽ lệch nhau không quá d giá trị.

Ở bước tính toán chi phí so trùng, mỗi 1 điểm ảnh bên trái được so sánh với d điểm ảnh bên phải nên sẽ có d giá trị chi phí so trùng tương ứng. Do đó, tạo thành ma trận 3 chiều gọi là Không gian chi phí so trùng (Matching Cost Volume)

Slide 10:

Bước 2, tích hợp, được thực hiện với mục đích tăng độ chính xác của giá trị đo lường đã tính được từ bước tính toán Matching cost.

Mỗi giá trị trong không gian chi phí so trùng được tính lại theo giá trị của nó cùng các điểm xung quanh.

Sau bước này, các giá trị ở không gian chi phí trở nên đáng tin cậy hơn và được sử dụng để tìm ra độ lệch ở bước 3.

Slide 11:

Như em đã trình bày ở bước tính chi phí so trùng, mỗi điểm ở ảnh trái được so sánh với d điểm ở ảnh phải để tính d giá trị chi phí so trùng.

Ở bước 3, tính toán độ lệch, từ d giá trị chi phí so trùng đã qua hiệu chỉnh, ta sẽ xác định điểm nào trong số d điểm được đem so sánh là tương ứng với điểm ảnh bên trái.

Kết quả của bước này là ma trận độ lệch của từng điểm ảnh.

Slide 12:

Ma trận độ lệch được tính ở bước 3 sẽ được hiệu chỉnh lại và trở thành kết quả của toàn bộ quá trình.

Trên hình ta có thể thấy kết quả đầu ra ứng với cặp ảnh đầu vào ở slide trước, những điểm có độ sâu nhỏ, tương ứng lệch lớn, sẽ có màu sáng, những điểm có độ sâu lớn, tương ứng độ lệch nhỏ, sẽ có màu tối.

Tiếp theo, em sẽ liệt kê nhanh 1 số cách tiếp cận cho từng bước.

Slide 13:

Ở bước tính chi phí so trùng, có thể chia cách tính chi phí thành các nhóm như sau:

* Các phương pháp tính toán chỉ dựa vào màu của cặp điểm: phương pháp này không đảm bảo kết quả tốt đối với nhiễu
* Các phương pháp tính toán dựa trên các giá trị màu xung quanh cặp điểm đang xét
* Các phương pháp biến đổi ảnh trước khi thực hiện tính toán
* Các phương pháp kết hợp nhiều hàm tính toán lại với nhau: đây là cách làm mới và đem lại hiệu quả cao.

Slide 14:

Ở bước tích hợp, dựa vào cách chọn các điểm hỗ trợ và trọng số hỗ trợ của chúng, có thể chia thành 2 nhóm chính:

* Nhóm không tương thích: các điểm nằm trên cửa sổ vuông lân cận điểm đang xét đều tham gia hỗ trợ và theo 1 phân bố chọn trước
* Nhóm tương thích: việc chọn các điểm hỗ trợ và mức hỗ trợ của chúng là riêng đối với từng điểm được xét, dựa trên những giả định riêng, đối với mỗi nghiên cứu.

Slide 15:

Disparity computation / optimization là bước tính toán ra giá trị độ lệch dựa trên giá chi phí so trùng giữa các cặp điểm ảnh có được qua hai bước tính toán chi phí và tích hợp.

Các phương pháp được sử dụng ở bước này được phân làm hai loại chính là cục bộ và toàn cục.

* Các phương pháp cục bộ chỉ tính toán độ lệch của một điểm xác định thông qua thông tin chi phí của chính điểm đó và xa hơn có thể là những điểm lân cận.
* Trong khi đó, các phương pháp toàn cục so trùng đưa ra giá trị được xem là tốt nhất trên toàn bộ bức ảnh

Phương pháp cục bộ cho kết quả kém hơn nhưng có độ phức tạp thấp.Ngược lại, các phương pháp toàn cục là bài toán tối ưu nên có chi phí tính toán cao và có chất lượng tốt.

Slide 16:

Mục tiêu của b ước Disparity refinement là hiệu chỉnh giá trị disparity đã tính toán đ ược, nhằm giảm thiểu các điểm ảnh lỗi cũng như xác định các vị trí ở vùng che khuất (chỉ nhìn thấy trên 1 tấm ảnh) để cho một kết quả đáng tin cậy hơn.

Quá trình hiệu chỉnh thường gồm 2 phần chính là

* Left-right validation: tìm các điểm không đáng tin cậy, và
* Disparity fill: hiệu chỉnh các giá trị lỗi

Slide 17:

Tiếp theo, em xin trình 1 số điểm chính trong các công trình nghiên cứu mà nhóm đã tìm hiểu. Để đảm bảo thời gian cho buổi báo cáo, em sẽ chỉ trình bày những nghiên cứu đề xuất được phương pháp cho kết quả tốt. Các nghiên cứu khác sẽ được đề cập ở phần sau.

Slide 18:

Phương pháp đầu tiên là từ 1 bài báo được công bố năm 2012 của tác giả Pérez và các cộng sự, dựa trên ý tưởng thực hiện bước lặp ở bước tích hợp chi phí cũng như hiệu chỉnh giá trị độ lệch.

Trong mỗi bước lặp, các giá trị độ lệch có mức tin cậy của chúng, và các giá trị chi phí so trùng cũng như độ lệch ở mỗi điểm ảnh được tính toán lại dựa trên giá trị chi phí ban đầu, độ lệch và mức tin cậy ở bước lặp trước.

Ngoài ra, phương pháp này còn sử dụng 1 giả định ở bước tích hợp, Các điểm khác về màu sắc và ở xa thì đóng góp nhỏ, từ đó đưa ra trọng số hỗ trợ tương thích.

So với chính nghiên cứu của nhóm tác giả được đề xuất 1 năm trước đó, nghiên cứu này đã có cải tiến trong việc tăng tốc quá trình xử lý bằng cách sử dụng những phương pháp đơn giản hơn nhưng vẫn đảm bảo kết quả tốt.

Slide 19:

Phương pháp thứ 2, là của nhóm tác giả Xing Mei, đề xuất năm 2011 với nhiều đóng góp như sau:

* Đề xuất cách tính chi phí so trùng dựa trên sự kết hợp 2 phương pháp dựa trên màu (là hàm AD theo khoảng cách Manhattan) và phân bố màu (là hàm biến đổi Census).
* Lựa chọn quá trình tích hợp tương thích đơn giản, nhờ đó các tác giả dễ dàng xây dựng vùng hỗ trợ tích hợp và việc tích hopwk được thực hiện với chi phí thấp
* Sau đó, độ lệch được điều chỉnh qua quá trình rất cẩn thận để phân loại và giảm các điểm lỗi nhiều nhất có thể.

Phương pháp này cho kết quả tốt ở nhiều bộ dữ liệu, tuy vậy, do sử dụng những giả định không phù hợp với đường giao thông nên phương pháp này cho kết quả không được tốt với dữ liệu ảnh mặt đường.

Slide 20:

Trong phương pháp thứ 3 được trình bày ở đây, nhóm tác giả L. T. Sach và các cộng sự đề cập năm 2009 với các đóng góp quan trọng nhất là

* Ý tưởng về việc xây dựng một bước tích hợp tương thích mạnh mẽ, đưa ra kết quả tốt với các trường hợp ảnh low texture, và
* Kĩ thuật lập trình để tăng tốc quá trình tính toán tích hợp.

Dựa trên giả định rằng những điểm ảnh nếu nằm trên cùng 1 vùng và có màu tương tự thường có giá trị độ lệch gần giống nhau, nhiều phương pháp stereo matching đã sử dụng kết quả phân đoạn ảnh theo màu để hướng dẫn quá trình tích hợp chi phí so trùng.

Do phân đoạn ảnh theo giá trị màu phức tạp và tốn chi phí tính toán, các tác giả xác định cạnh của ảnh đầu vào, với mục đích hướng dẫn quá trình tích hợp chi phí. Dựa vào thông tin về canh, các điểm thuộc vùng sẽ có các trọng số khác nhau.

Nhờ đềra các giả định hợp, phương pháp này đưa ra kết quả tốt cho trường hợp ảnh lowtexture.

Slide 21:

Ngoài ra, trong quá trình thực hiện luận văn, nhóm còn tìm hiểu nhiều công trình nghiên cứu khác, những nghiên cứu này đều được trình bày chi tiết trong luận văn.

Bây giờ em sẽ trình bày về phần công việc trọng tâm được hiện thực trong luận văn. (đến slide 31, bỏ qua các slide 22-30)

Slide 31:

Nội dung của mục này gồm có 2 phần được phân chia rõ, phần thứ 1 là thiết kế và xây dựng benchmark cho các công trình nghiên cứu về Stereo Matching và phần thứ 2 là đề xuất và hiện thực phương pháp ứng dụng kết quả độ sâu tính được từ quá trình Stereo Matching vào bài toán phát hiện đối tượng trên đường.

Slide 32:

Trước hết, em xin trình bày nội dung của phần thiết kế và xây dựng benchmark.

**Công việc đầu tiên và rất quan trọng là xác định được những công trình nghiên cứu về Stereo Matching để khảo sát, tìm hiểu.**

Một may mắn cho nhóm khi thực hiện đề tài, đó là cộng đồng Stereo Matching đã phát triển mạnh mẽ trong khoảng 20 năm trở lại đây và xây dựng Middlebury, 1 website chuyên liệt kê và đánh giá các phương pháp, cũng như cung cấp các tập dữ liệu đa dạng để tiến hành thực nghiệm.

Nhóm chọn những phương pháp được xếp hạng cao trong bảng đánh giá của Middlebury để khảo sát dựa trên những tiên chí của nhóm.

**Công việc thứ 2, là lựa chọn tiêu chí để phân loại các phương pháp Stereo Matching đã được lựa chọn.**

Trong quá trình tìm hiểu mỗi công trình nghiên cứu, bọn em sẽ phân tích ý tưởng, cách thức thực hiện của tác giả cho từng bước của mỗi nghiên cứu. Từ phân tích của mình và hình thức thiết kế bảng phân loại giống như trong báo cáo đánh giá của Scharstein và Szeliski, bọn em đã xây dựng bảng phân loại cho riêng mình, sẽ được giới thiệu trong Slide sau.

Bọn em cũng tập trung gom các phương pháp thành các nhóm, những phương pháp trong cùng 1 nhóm cùng chú trọng vào việc cải tiến 1 bước cụ thể trong quá trình Stereo Matching cơ bản. Có những nghiên cứu cải tiến nhiều - có thể 2, 3 hoặc cả 4 bước - bọn em sẽ đánh giá xem bước cải tiến nào là đóng góp quan trọng nhất của tác giả, để từ đó đưa vào nhóm phù hợp.

**Công việc tiếp theo sau khi đã khảo sát các phương pháp là hiện thực lại và thực nghiệm để đánh giá kết quả giải thuật, cũng như kiểm nghiệm lại những hiểu biết của nhóm.**

Vì những tập dữ liệu trên trang Middlebury cùng chuẩn vàng của chúng đã được sử dụng rộng rãi để thực nghiệm, nên nhóm em cũng sử dụng lại, và kết quả thực nghiệm đã được trình bày trong luận văn, cũng như phần cuối của bài báo cáo này.

Slide 33:

Sau quá trình thiết kế và xây dựng benchmark, nhóm em đã có bảng phân loại như sau:

* Các phương pháp được gọi nhóm theo tên tác giả
* Ở 4 bước của mỗi phương pháp, bọn em sẽ phân tích các ý tưởng đề xuất chung nhất mà tác giả đóng góp. Một số tác giả chỉ đề cập đúng 1 số phần mà họ quan tâm sâu sắc, thì những phần còn lại sẽ được gán nhãn “Không đề cập”
* Trong bảng cũng thể hiện rõ các phương pháp đã được gom thành 4 nhóm, nhóm 1 gồm các phương pháp cải tiến quá trình tích hợp, nhóm 2 gồm các phương pháp tính toán độ lệch dựa vào phương thức toàn cục, nhóm 3 gồm các phương pháp tính toán độ lệch dựa vào phương thức cục bộ nhưng có bước hiệu chỉnh để làm tốt kết quả, nhóm 4 gồm các phương pháp đưa ra những giả định dựa vào kinh nghiệm và đạt được kết quả tốt.

Sau khi hiện thực benchmark đối với các phương pháp Stereo matching, nhóm thực hiện phần thứ 2 là đề xuất và hiện thực phương pháp ứng dụng kết quả độ sâu tính được từ quá trình Stereo Matching vào bài toán phát hiện đối tượng.

Slide 34:

Em xin trình bày nội dung phần thứ 2.

Do giới hạn của đề tài là phát hiện đối tượng trong môi trường giao thông, do đó, nội dung công việc của phần này là ứng dụng kết quả độ sâu tính được từ giải thuật Stereo Matching vào bài toán phát hiện đối tượng tham gia giao thông trong cặp ảnh mặt đường.

Quá trình ứng dụng gồm có 2 bước, bước thứ nhất là ước lượng độ lệch của cặp ảnh, dựa vào 1 trong các phương pháp đã khảo sát. Bước thứ 2, từ độ lệch, ước lượng độ sâu và áp dụng vào 1 phương pháp phát hiện đối tượng để tìm ra mặt đường và các đối tượng trên mặt đường.

Do đặc điểm low texture của ảnh giao thông, nhóm đã đánh giá và quyết định dựa vào phương pháp xác định độ lệch do tác giả L.T.Sách và các cộng sự đã công bố trong nghiên cứu của họ để đề xuất giải thuật của mình.

Slide 35:

Trong hình trên, sơ đồ của giải thuật tính độ lệch của nhóm sử dụng gồm các phần chính như sau:

* Tính toán chi phí và tính toán độ lệch sử dụng những phương pháp đơn giản
* Tích hợp chi phí là phần quan trọng nhất, đóng vai trò quyết định đến chất lượng của đầu ra.
* Quá trình xác định cạnh ở cặp ảnh đầu vào có vai trò hỗ trợ đến bước tích hợp chi phí, đây cũng là 1 điểm nổi bật của giải thuật.

Chúng ta sẽ đi vào chi tiết của giải thuật tính độ lệch ở các Slide tiếp theo.

Slide 36:

Ở bước tính toán chi phí, phương pháp SAD, 1 phương pháp sử dụng cửa sổ hỗ trợ, được lựa chọn sử dụng vì nó có ít chịu ảnh hưởng của nhiễu, có đem lại kết quả đáng tin cậy với thời gian tính toán không quá cao.

Trọng tâm của giải thuật là bước tích hợp chi phí. Bước này được tiến hành theo 2 giai đoạn là tích hợp theo chiều ngang và tích hợp theo chiều dọc.

Nhóm áp dụng 3 giả định về ảnh giao thông đã được đề cập trong nghiên cứu của nhóm thầy L.T.Sách và cộng sự.

* Thứ nhất, những điểm ảnh nằm trên cùng 1 vùng và có màu tương tự thường có giá trị độ lệch gần giống nhau.
* Thứ nhì, do đặc điểm của ảnh thu được từ góc nhìn của phương tiện giao thông, độ lệch của các điểm ảnh nằm trên cùng hàng ngang thường ít thay đối, trong khi độ lệch thay đổi nhiều với các điểm ảnh nằm trên cùng hàng dọc.

Do đó, nhóm dùng phương pháp nhân mặt nạ LoG và lấy ngưỡng để có được mặt nạ cạnh nhị phân, sử dụng để hướng dẫn quá trình tích hợp, cụ thể là xác định trọng số cho các điểm hỗ trợ.

Ở giai đoạn tích hợp theo chiều ngang, việc tích hợp được tiến hành theo từng hàng ngang với của sổ tích hợp rất dài trong khi tích hợp trên từng dọc với cửa sổ tích hợp rất ngắn.

Việc tiến hành quá trình tích hợp trên 1 hàng sẽ được mô tả trong slide tiếp theo.

Slide 37:

Giả sử ta sẽ tích hợp các giá trị của không gian chi phí, tại lát cắt U-V ứng với giá trị d = dj và tại dòng ứng với giá trị v = vi

Các điểm ảnh hỗ trợ nằm trong cửa sổ tích hợp và nhận các giá trị lần lượt là từ a, b đến g, còn điểm p là điểm cần tính lại giá trị chi phí

Bên dưới tương ứng là giá trị mặt nạ cạnh tương ứng của các điểm trong cửa sổ tích hợp. Các điểm có giá trị giống nhau cùng thuộc về 1 phân đoạn.

Dễ dàng thấy, bên trong vùng cửa sổ, các điểm ảnh thuộc về 3 phân đoạn khác nhau, trong đó điểm p thuộc phân đoạn thứ 2.

Những điểm nằm trong phân đoạn thứ 2 có trọng số hỗ trợ là 1, còn những điểm thuộc các phân đoạn còn lại có trọng số là alpha.

Chi phí tích hợp mới của điểm p được tính bằng tổng các trung bình cộng của các phân đoạn sau khi đã nhân trọng số hỗ trợ.

Sau khi tíh xong tại điểm p, quá trình tiếp tục với điểm kế bên trên cùng dòng và thực hiện tuần tự trên từng dòng, theo chiều ngang rồi sau đó là chiều dọc.

Slide 38:

Từ không gian chi phí đã được tích hợp, quá trình tính toán độ lệch tại mỗi điểm ảnh được tiến hành cục bộ, ở đó, điểm được chọn sẽ tương ứng với giá trị chi phí thấp nhất.

Giá trị hiệu chỉnh của giải thuật do nhóm đề xuất chỉ nhằm thực hiện 1 thao tác, đó là kiểm tra và loại bỏ những điểm không hợp lệ khỏi quá trình tính toán tiếp theo.

Đến đây thì giải thuật tính toán độ lệch đã hoàn tất, kết quả nhận được là ma trận độ lệch đã qua hiệu chỉnh.

Slide 39 - 40:

Mỗi giá trị d của ma trận độ lệch hiệu chỉnh chính là thông tin độ sâu của của điểm ảnh tương ứng. Do đó có thể xem ma trận độ lệch là toàn bộ thông tin độ sâu của ảnh trong không gian.

Chiếu các điểm ảnh lên mặt phẳng Oyz và định tập điểm mẫu, là điểm thấp nhất tại mỗi giá trị độ sâu và có tần suất lớn hơn ngưỡng cho trước.

Các điểm mẫu này là các điểm được xem là thuộc về mặt đường. Tập điểm mẫu sẽ được đưa vào biến đổi Hough để tìm mặt đường, được ước lượng bằng tập các đoạn thẳng.

Điểm biên của vùng mặt đường trống được xác định bằng phương pháp cục bộ WTA.

Sau khi ước lượng được mặt đường, các đối tượng trên mặt đường được xác định bằng tọa độ không gian so với mặt đường ước lượng bằng cách so sánh và lấy ngưỡng.

Qua các Slide này, em đã trình bày xong toàn bộ phương pháp do nhóm đề xuất để áp dụng Stereo Matching vào phát hiện đối tượng trên đường.

Slide 41:

Phần tiếp theo trong nội dung thực hiện mà nhóm đã đề xuất là việc thực nghiệm các phương pháp Stereo Matching đã tìm hiểu, cũng như phương pháp mà nhóm đã đề xuất và hiện thực.