Usability of a Domain-Specific Language for a Gesture-Driven IDE

Michaela Bačíková, Martin Maričák, Matej Vancík
Technical university of Košice
Letná 9, 042-00 Košice, Slovakia
Email: michaela.bacikova@tuke.sk, {martin.maricak, matej.vancik}@student.tuke.sk

Abstract—User interfaces (UIs) are advancing in every direction. The usage of touch screen devices and adaptation their UIs lives its boom. However integrated development environments (IDEs) that are used to develop the same UIs are overseeping the time. They are directed to developing usable software, but forgot to be usable by themselves. Our goal is to design a new way of user interaction for common IDEs with the help of touch. The target group are hybrid devices formed by a physical keyboard and either an integrated, or separate, touch screen display. In this paper we describe a set of general purpose and domain-specific gestures which represents a language for working with a touch-driven IDE and provide a method their design. We performed two studies with developers from industry and university and developed a prototype of a gesture-driven IDE to evaluate the usability of the presented approach.

I. INTRODUCTION

The accessibility of alternative devices such as wearable gadgets, smartphones, tablets, tablet PCs, wall-sized displays, touch-enabled displays, kinect and others, has given a rise to different design guidelines for application user interfaces (UIs) which, though still young and imperfect, become more and more dominant. According to the world statistics a large part of interaction has moved to touch devices and the vast majority of mobile devices now supports touch control. Touch interfaces are natural and easy to get used to and they also affect the way people try to control other devices such as common PCs or notebooks. Therefore, hybrid devices1 have emerged and operation systems try to adapt their UIs and functionality to touch.

Although touch control support has moved a large leap forward, it still has its disadvantages. Touch gestures are not entirely uniform on all platforms and in many cases, a touch display alone is not sufficient to fully handle all desired functionalities. A typical example is writing on a virtual keyboard. Writing on a virtual keyboard is slow, imprecise and the keyboard occupies a large part of the screen [1]. Still, the situation is much better than before and the rise of platform guidelines for touch devices with different screen sizes helps the situation.

Despite the disadvantages, the situation is much better than in case of integrated development environments (IDEs). They are made to improve developer productivity as much as possible [2] but it is as if developers were primarily focusing on making UIs and interaction better just for their users while forgetting about themselves. The support of touch for the most common IDEs literally overslept the time. Not only the current IDEs are not ready for advanced work on touch displays, they are not even prepared for everyday touch use stereotypes and are largely limited to interaction via traditional hardware devices (e.g., the ubiquitous keyboard and mouse) [3]. It is not possible to use swiping to scroll, pinch to zoom or tapping the same way as in every other touch-enabled UI. From the IDEs we have analyzed, the only IDE ready touch is Eclipse, statistically the most used IDE for Java language. However, it still offers just fundamental features. As for the second line in line, IntelliJ IDEA, despite marked innovative and usable, does not even support the most fundamental ones. Instead of browsing or scrolling through code, the swipe gesture selects text, infringing the common interaction stereotypes and adjustment of panel sizes or zooming by using pinch gesture is not possible at all.

Programming requires the speed of interaction and creativity, touch support is unambiguously a step forward because it brings a different form of user experience and according to Greene [4], supports creative thinking. Our goal is not to force the work exclusively via touch gestures but to create a symbiosis of keyboard, touch and, alternatively, mouse. Begel [5] experimentally verified that programmers are so used to physical keyboard that it would be disadvantageous or at most impossible for them to explicitly use virtual keyboard, which, in addition, has a low precision and does not enable writing by ten fingers.

Supporting common general gestures in IDE may not be enough to speed up the developer work. A specific solution is necessary to not just support the basic gestures, but enable advanced interaction adjusted for a specific IDE platform and programming language. In this paper, we show a method of designing a set of gestures for such advanced interaction as a way of designing the lexical part of a visual language. The focus is mainly on usability [6], [7], [8], domain usability and specificity [9], simplicity of use, learnability and compliance with standard design patterns and guidelines.

The main contributions of this paper are:

1. Definition of more complex general-purpose gestures in the context of IDEs.

---

1In this paper, the term hybrid devices will be used to refer to devices with a keyboard and a touch display such as tablet PCs or combinations of desktop PCs/notebooks with an external touch display.
A method of collecting, identifying and designing a set of domain-specific gestures via multiple user surveys and interviews including industry programmers and gesture recorder.

Verbal definition of a domain-specific language of gesture patterns to support highly specialized work with code and consequently a design and realization of drawn gesture input and recognition in a form of a gesture-driven IDE for Android. When compared to a physical keyboard, we hypothesize that touch gestures are more learnable and memorable than key shortcuts.

Evaluation of the proposed approach by means of usability evaluation.

Although our prototype is designed for Android devices, we do not focus on touch devices explicitly (although they may be a possible target), our target group are hybrid devices.

II. A TOUCH-ENABLED IDE

Here, we will describe a design of fundamental gesture interaction for touch devices related to IDEs. Elementary gestures are based on usability guidelines [10], [11], analysis of existing desktop and touch-based IDEs and discussions with industrial developers. Primarily, we followed the Google Material design guidelines [12] that also define the most common gesture patterns (Fig. 1).

Breaking the stereotypes of GPGs can cause serious usability issues, mainly in the case of advanced or experienced touch device users. Therefore it is good to abide the standard interaction in case of this category. In case of an IDE, it is advantageous to design additional gestures similar to general-purpose ones by using multiple fingers to support more action shortcuts. When designing the set of GPGs, we primarily followed the guidelines of the Android platform and we were also inspired by the iOS platform guidelines for multi-touch gestures [14].

Based on the simple gestures displayed in Fig. 1, we further designed the following additional multi-touch gestures for common IDEs:

1) Double-touch to select words.
2) Tripple-touch to select lines.
3) Two-finger horizontal swipe for undo/redo inspired by the usage on Apple trackpads [14].

III. DOMAIN-SPECIFIC LANGUAGE FOR A GESTURE-DRIVEN IDE

IDEs are used in the domain of programming and to be able to support at least the standard interaction with an IDE, we need to design domain-specific gestures.

The set of gestures supported by an application can be perceived as a sign language, which the user has to learn to be able to communicate effectively. A correct conceptual model helps users to better understand actions performed upon a UI, thus speeds up the learning process.

From this point of view, the list of domain-specific gestures (both general-purpose and drawn) described further in this paper can be perceived as a domain-specific language of the touch-enabled IDE:

1) Concrete syntax is defined by the gestures themselves: touch or movement, finger count, touch count and movement directions.
2) Abstract syntax is defined by the actions that the gestures cause.
3) Semantics is defined by their type and implementation.

Because of the limited space and for we design a simple visual language (the set of gestures is kept small), we will not formally describe the semantics of the gesture DSL in this paper. We expect that common language constructs, such as loops, conditions, comments or class and method declarations are known to common programming language designers. Therefore we expect the semantics to be clear from the purpose of the gesture. For each gesture, the semantics is a generated programming language construct. We use templates to generate the constructs similar to the ones used in common IDEs using the Velocity template engine. Concrete and abstract syntaxes are described verbally.

A. General-Purpose Gestures

We designed the following GPGs for touch-enabled IDEs:

1) Swiping:
   a) Two-finger swipe for "stuck shift" - gradual shift from method to method by swiping two fingers.
b) Three-finger swipe to the end/beginning of the file.

2) Code folding:
   a) Three-finger pinch for method folding.
   b) Four-finger pinch for "semantic zoom".
   c) Five-finger pinch to fold all class members.

"Semantic zooming"\(^2\) is a way of grouping content into specific predefined categories. In case of an IDE it would be a combination of related methods or blocks of code such as getters, setters, void methods, public methods, imports, nested classes, etc., into a single group. After performing the gesture, the user is prompted to select a semantic group. After selection, everything else not belonging to the selected group is folded. This enables to focus only on the important parts of the code.

In case of pinch gestures, the action of pinching-in standardly folds class members and the opposite direction of pinching-out causes unfolding action.

B. Drawn Gestures

As we stated in section II, drawn gestures are not subject to any specifications or conventions. Therefore their design needs to be approached with caution. They are specific for the target domain of use, therefore domain experts should be included in the design and development process. A domain analysis was performed to determine the way how programmers interpret different language constructs visually. We interviewed multiple developers from industry and their claims were supplemented by a survey. In the next subsections we will describe the results of the survey that was performed in two iterations.

1) First Survey - Collection of Patterns: In order to elicit the correct conceptual model most effectively, a support application for Android touch devices called "Gesture Recorder" (Fig. 2) was created. The whole survey process was performed in the application, using a Google Nexus 7 tablet, which asked users to interpret given programming language constructs by directly inducing gestures on touch screen. We used a standard Android tablet for the survey.

We wanted to cover a wide spectrum of programmer experience, therefore we included 68 participants, from which 8 considered themselves advanced programmers, 38 intermediate and 22 were beginner programmers. During the survey, each participant performed the survey independently to prevent influence and was assisted by one member of our research team in order to be able to correctly work with the survey application.

The survey application prompted the participants to interpret the following constructs by a drawn gesture: class declaration, comment, variable declaration, loop statements (both for, while and do-while), condition statement, method declaration and deleting a line.

For each prompt, an example of a language construct was displayed in the application to help the user to better connect his/her mental model with a new gesture. After all language constructs, participants were given a possibility to add their own idea for a new gesture and language construct. Then they were to fill a questionnaire about their programming experience. The results for each participant were sent to a distant server, manually checked and analyzed.

   a) Evaluation Method: Similarity of gesture shapes was taken into account via their abstractions into patterns. Differences in size, shape and format were abstracted into a single unified pattern while the direction of drawing was preserved. Differences were neglected only if the conceptual model of the gesture interpretation remained unchanged. The patterns with low repetition rate were omitted. Fig. 3 shows the simplified survey results. Each group represents one language construct and indicates the most common gesture pattern and the number of participants that used the pattern. The direction of drawing is indicated by the small circle representing the final point of drawing.

   b) Results: The results in the case of a class were influenced by the fact that the respondents come from Slovakia, where class is called "trieda" and begins with the letter "T". Many of the participants also included "N" as in "new", but the number was not as significant. Both were mainly beginners or at least advanced programmers and from our experience, sooner or later older and more experienced programmers begin to in English language explicitly and do not perceive the class declaration and instantiation as the same concept anymore. In case of variable declaration, the results were unclear, therefore we decided not to include the gesture without further research. In case of cycles, as for the shape, the results jointly indicate a circle, as for the direction, the results are unclear.

2) Second Survey - Suitability of Collected Patterns: Because the results from the first survey were very variable, we decided to perform a second survey. The goal was to determine which shape of the previously identified ones was the most appropriate for the given language construct. The second survey was performed on 65 new participants. In the first survey we covered different levels of programmer experience to get universal results but this decision has proven to be wrong. In the second round we rather decided to target

\(^2\)More information about semantic zooming can be found in Microsoft official guidelines: https://msdn.microsoft.com/en-us/library/windows/apps/hh465319.aspx
As for the final question, most of the respondents agreed on a horizontal line for line deletion.

IV. IMPLEMENTATION OF THE GESTURE DSL

At this point we will describe one of the ways of implementing recognition of the described gestures. We chose Android as a model platform mainly because of its accessibility, openness and good support of developing and recognizing custom gestures. However, the mechanics of gestures and the described procedures are common for any platform. The goal is not to point out to the details but to show the method of developing a gesture recognition engine.

We developed two prototypes of a simple gesture-driven IDE with the goal to design the gesture recognition algorithms and tested it with industrial developers. We focused on maximal usability. We redesigned and implemented recognition of both complex GPGs and drawn gestures.

Recognition of GPGs is implemented via native platform libraries. Recognition of drawn gestures is more complex and uses machine learning and recognition probabilities. Due to the limited space, we will only describe the implementation of domain-specific drawn gestures.

A. Recognition of Drawn Gestures

To recognize drawn gestures we used the Android com.android.gesture package containing classes such as Gesture, GestureOverlayView and Prediction enabling possibilities to record, store, read and recognize gestures. The package is undocumented, therefore we will describe the design in more detail. GestureOverlayView is a graphical component that enables the actions of gesture recording and displaying it as curves. The recorded and predefined gestures are stored in a binary file located in application memory and contain a list of points and a single name may be assigned to each gesture.

Recognition works on the principle of comparing a drawn gesture with the whole file content and the result is a list of identified patterns with assigned probabilities. However, the default recognition is not very accurate, especially if the stored patterns are slightly similar to each other or if the drawn gesture did not imitate the target pattern too precisely. The more patterns are stored in the database, the less precise is the recognition. In case there are multiple gesture patterns, the recognition rate can be less than 40% which is unacceptable. Taking this fact into account, we needed to adjust the default recognition process so that each pattern could have a database of multiple possible shapes. We recommend to create the databases by recording pattern shapes from different users and multiple times and also to record the gestures drawn by the IDE users during their first use by means of a tutorial.

The recognition probability \( p \) of match is determined by a positive number. The exact meaning of \( p \) was not clear since the package is documented, thus the value was determined empirically for each pattern as stated below. In general, the gesture is successfully recognized if \( p > 1 \) and the highest probability is considered as correctly recognized. This applies only after verifying the number of traces forming the gesture. If a circle is recognized after drawing two mirror-inverted \( C \) characters, the result will not be considered correct, even though it has the highest \( p \) value.

The database for each pattern was created by drawing multiple slightly deformed shapes and in both drawing directions. Because of significant differences between probabilities for patterns with different complexity, an optimal \( p \) interval was determined empirically for each pattern as follows:

<table>
<thead>
<tr>
<th>Class declaration</th>
<th>Variable declaration</th>
<th>Comment</th>
<th>Do-while loop</th>
</tr>
</thead>
<tbody>
<tr>
<td>13</td>
<td>7</td>
<td>35</td>
<td>16</td>
</tr>
<tr>
<td>4</td>
<td>8</td>
<td>10</td>
<td>6</td>
</tr>
<tr>
<td>6</td>
<td>56</td>
<td>23 other</td>
<td>6</td>
</tr>
<tr>
<td>7</td>
<td>4</td>
<td>2 other</td>
<td>2</td>
</tr>
<tr>
<td>38 other</td>
<td></td>
<td></td>
<td>38 other</td>
</tr>
</tbody>
</table>

While / for loop | Condition | Method declaration | Line deletion |
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>16/12</td>
<td>9</td>
<td>17</td>
<td>10</td>
</tr>
<tr>
<td>9/16</td>
<td>8</td>
<td>51 other</td>
<td>9</td>
</tr>
<tr>
<td>43/40 other</td>
<td>3</td>
<td>49 other</td>
<td></td>
</tr>
</tbody>
</table>

Fig. 3. The results of the first survey
Loop gesture - is considered to be recognized if \( p > 4 \).
Condition gesture - the most complex shape of all because of two changes in the direction. Recognized when \( p > 3 \).
Class gesture - a higher value has proven to be optimal. This gesture is recognized if \( p \geq 5 \).
Comment gesture - a simple gesture of two consequent lines. Recognized if \( p > 10 \)
Line deletion gesture - the most simple gesture, almost identical to comment, but with only one line and different direction. Valid recognition if \( p \geq 10 \).

The values were determined by re-entering gestures multiple times and observing the \( p \) value after recognition. We verified both positive and negative cases and with the stated \( p \) values, we achieved almost 100% success rate.

Since the recognition database is built using a simple form of machine learning and the recognition thresholds are determined empirically by applying probabilities, it is important to consider the number of available patterns in the database. Low number of predefined gestures that are not very complex and similar to each other is optimal. Also, it is very important not to continue the machine learning process during application use, because the user might draw incorrect shapes and hinder the recognition of the predefined gestures. Instead, we recommend to allow the user to create his/her own gestures and teach the correct and incorrect shapes to the recognition algorithm. This is also a problem of many systems for recognizing handwritten text. Each person has his/her own writing style and when supporting as many characters as the whole alphabet, the recognition can be incorrect multiple times and lead to reluctance and rejection of using the system. Moreover, the user is only able to remember a limited number of gesture shortcuts and forcing his/her to remember a whole database of patterns can reduce usability.

V. GESTURE DSL - USABILITY TEST

To evaluate the gesture DSL, we created a prototype of a simple gesture-driven IDE for Android according to the design described in this paper and performed a usability evaluation.

Our target group were Java programmers, experienced in using IDE and minimally in using a touch-enabled smartphone or a tablet. For the reasons indicated by Nielsen [16], [7], we performed the usability evaluation with five participants and we targeted at qualitative evaluation. The participants were given tasks related to tutorial, GPGs and 2 tasks targeted at general work with the IDE.

Each testing was performed in the presence of one member of the research team. The sample was represented by five experienced developers with 1-5 years of practice in industry, using mostly IntelliJ IDEA. To avoid influence, the test was performed separately with each participant. In all cases we used Google Nexus 7 tablet and each time the application was reset to its original settings. Two of the respondents were given an external keyboard.

At the beginning, the participants were briefly familiarized with the research and with the testing process and they were given an in-app introductory tutorial. After finishing the tutorial, the participants were to perform the tasks on a sample class and to express their opinion.

A. Observations and Conclusions

Gesture discoverability was 100%, each participant was able to perform the tasks successfully. The feedback received from the subjects indicated that almost all participants liked the idea of manipulating with the editor using GPGs and generating parts of code using drawn gestures. However, the idea of using it on a tablet was not accepted, each participant expressed the wish to use the gesture language on a common IDE which indicates the potential of our approach on hybrid devices. We concluded that the negative opinion towards tablets is related to writing code using a virtual keyboard. The results were definitely better in case of participants that used an external keyboard.

The gesture for folding all methods was marked most useful, however the users expressed the idea of subsequent use of the gesture multiple times to fold methods to regions. All of them liked the semantic zoom feature and indicated that it could replace the common IDE feature of code structure preview. They suggested to add a filter for constructors, member variables and a filter for displaying all method names.

The possibility to define custom gestures was described as a feature with high potential. Each developer mentioned an example of use, such as definition of custom predefined code parts and templates or combining with macros.

The design of new gestures is a major challenge. The gestures have to be simple, easy to learn and quick to write. This holds for the set presented in this paper, but with a growing number of gestures, their complexity might rise significantly, which could complicate their use. It might be beneficial to explore multiple combinations of gestures and patterns [17] [18] of their use.

VI. RELATED WORK

Biegel et al. [19] propose an inspirational approach to "touchifying" an IDE. They highlight multiple issues of common IDEs related to supporting touch handling in general. Authors present a solution on Eclipse IDE. Compared to our gesture set, they used already existing GPGs to support multiple functionalities. In their study, they did not include domain-specific drawn gestures. The fact that programmers rather prefer the combination of a traditional PC, touch display and a mouse shown in their paper supported our findings and confirms that this approach has a potential to be used on hybrid devices.

There are multiple works that focus on code refactoring triggered by gestures, such as the works by Murphy-Hill et al. [20] and Raab et al. [21]. Compared to their solution, we use a small set of gestures for commonly used features and for the sake of usability we never use multiple composed gestures to trigger a single operation. Lee et al. [22] propose a solution where refactoring can be triggered via drag-and-drop gestures but in contrast to our work they optimized the gestures for using a mouse.
Hesenius et al. [23] introduced an environment for tablets with advanced features but focused primarily on creating a new environment, not adaptation to hybrid devices. Delimarschi et al. [3] describe a natural UI for graphical IDE using a Kinect voice and physical gesture commands. Although we favour the potential of voice commands as means to speeding-up the developer work, voice recognition is still not mature enough to fully support multiple languages. As for physical gestures, the potential is less definite. Edwards and Barnette [24] unsuccessfully used tablet PCs with a pen in a laboratory programming course without adapting the software to this novel device.

Begel et al. [5] mentioned that it is possible that developers are less efficient by using a natural interface instead of typing. This is related to the fact that typing on a physical keyboard is already ingrained in every developer’s habits, while more natural interaction might not be as standard in the development process. We agree that in the case of IDEs, not each natural interaction type is beneficial, however we argue for touch input being advantageous based on our observations and user feedback. Direct manipulation characteristics of touch helps to reduce the cognitive load [19]. Further investigation is needed to support this claim.

VII. Conclusion

We introduced the concept of a DSL of gestures in the context of IDEs. We performed multiple studies to explore the visual representations of conceptual models of programming constructs and developed a prototype tool which served to conduct a pilot study to assess the feasibility of our approach. Despite the challenges and current limitations, we believe that it has the potential to improve and speed-up development and ease of use. To make the approach more accessible and verifiable in the context of hybrid devices, we have set forth the following goals: 1. Implement a prototype into a common desktop IDE such as Eclipse, IntelliJ IDEA or NetBeans in form of a plug-in. 2. Integrate learning capabilities that would allow the system to adapt to the user and his/her drawing style and allow the user to add new custom gestures via the existing IDE-native code template engine. 3. Generalize the approach beyond Android and Java, for a variety of operating systems and programming languages. 4. Experimentally identify the optimal number, complexity and combinations of drawn gestures.

Acknowledgment

This work was supported by VEGA Grant No. 1/0341/13 Principles and methods of automated abstraction of computer languages and software development based on the semantic enrichment caused by communication.

References